Proceedings of the Indian Academy of Sciences

Vol. 91, January–December 1982

CONTENTS (Mathematical Sciences)

Proc. Indian Acad. Sci. (Math. Sci.), Vol. 91, Number 1, January 1982, pp. 1-15.

Ramanujan and Dirichlet series with Euler products

S S RANGACHARI
School of Mathematics, Tata Institute of Fundamental Research, Homi Bhabha Road, Bombay 400 005, India.

MS received 10 March 1981 ; revised 7 July 1981

Abstract. Ramanujan, in his unpublished manuscripts, had written down without proof, explicit linear combinations of cusp forms whose Mellin transforms possess Euler products in the sense of Hecke. All these results are proved here and their connection with the work of Hecke, Rankin and Serre is pointed out.

Keywords. Dirichlet series ; Euler products ; Ramanujan's work.

1. Introduction

In addition to the different results on Dirichlet series with Euler products in his famous paper [7], Ramanujan had stated some more results in the manuscripts recently discovered [1, 8]. We shall discuss in this paper all the Dirichlet series with Euler products found in his published and unpublished work. We divide them for the sake of convenience into three parts although there are some overlaps. We shall remark at the end the connections between the different Euler products.

(i) Euler Products in [7]—they are all Dirichlet series attached to $\eta^k\left(\frac{24z}{k}\right)$, $k|24$ (η: Dedekind's η—function) or Dirichlet series attached to cusp forms occurring in "sums of squares" problem.

(ii) Euler Products in [8]—these are Dirichlet series attached to cusp forms of the type $(\eta(z)\,\eta(Nz))^k$, $k|24$ and $(N+1)|24$ which occur in congruence problems.

(iii) Euler Products in [1]—these are Dirichlet series attached to cusp forms on cycloidal subgroups of the modular group and presumably studied by Ramanujan for congruence properties for $p(n)$. (See appendix in [6]).

The results in (1) were proved by Mordell in [4] and in a more general fashion by Hecke in (35, 36 in [2]) the final result being the proof by Deligne of the Ramanujan conjecture. Some of the results in (2) were proved by Hecke in ((41) in [2]) who was obviously unaware of Ramanujan's statements.

The results in (3) have not been proved so far, although a few of them belong to the category (1) and hence proved by Mordell and Hecke. We shall supply

proofs of all the results in (3) and also fill in the gaps left by Ramanujan. For the same, we need the results of Rankin [9].

The most remarkable thing about the results in (3) is the fact that certain linear combinations of Dirichlet series have Euler products, the first published examples of which we encounter only in Hecke, as observed by Birch in [1].

2. Ramanujan's results on Dirichlet series with Euler product

We shall list below all the Euler product developments given by Ramanujan, both published and unpublished. The first list is from his paper [7] and the second list is from his unpublished papers [8] and the third list as given in Birch [1].

(*a*) Euler product development for Mellin transforms of

$$\eta^{a}\left(\frac{24z}{a}\right) \text{ with } a \mid 24 \text{ and for Dirichlet series } \sum_{n=1}^{\infty} e_k(n) \cdot n^{-s}$$

($k = 10$, 12, 16) as given in [7].

(*b*) Euler product development for the Mellin transforms of

(1) $\eta(z)\,\eta(7z)$, (2) $\eta(z)\,\eta(11z)$, (3) $\eta(z)\,\eta(23z)$, (4) $(\eta(z)\,\eta(11z))^2$, (5) $(\eta(z)\,\eta(7z))^3$.

More explicitly, they are as follows : (set $x = \exp.(2\pi i z)$)

(1) Let $\sum\limits_{n=1}^{\infty} \phi(n)\, x^{n/3} = x^{1/3} \prod\limits_{n=1}^{\infty} (1 - x^n)(1 - x^{7n})$. Then

$$\sum_{n=1}^{\infty} \phi(n) \cdot n^{-s} = (1 + 7^{-s})^{-1} \prod_p (1 + p^{-2s})^{-1} \prod_q (1 - q^{-2s})^{-1}$$

$$\prod_r (1 + r^{-s})^{-2} \prod_t (1 - t^{-s})^{-2}$$

where $p \equiv 2, 8, 11 \pmod{21}$, $q \equiv 5, 17, 20 \pmod{21}$

$r \equiv 1, 4, 16 \pmod{21}$ which are of the form $9A^2 + 7B^2$

$t \equiv 1, 4, 16 \pmod{21}$ which are of the form $A^2 + 63B^2$

(2) Let $\sum\limits_{n=1}^{\infty} \phi(n) \cdot x^{n/2} = x^{1/2} \prod\limits_{n=1}^{\infty} (1 - x^n)(1 - x^{11n})$. Then

$$\sum_{n=1}^{\infty} \phi(n) \cdot n^{-s} = (1 - 11^{-s})^{-1} \prod_p (1 - p^{-2s})^{-1} \prod_q (1 + q^{-s} + q^{-2s})^{-1}$$

$$\prod_r (1 - r^{-s})^{-2}$$

where $p \equiv$ quadratic non-residue mod 11

$q \equiv$ quadratic residue mod 11 of the form $11A^2 + B^2$

$r \equiv$ quadratic residue mod 11 not of the form $11A^2 + B^2$

(3) Let $\sum_{n=1}^{\infty} \phi(n) \cdot x^n = x \prod_{n=1}^{\infty} (1 - x^n)(1 - x^{23n})$. Then

$$\sum_{n=1}^{\infty} \phi(n) \cdot n^{-s} = \prod_p (1 - p^{-2s})^{-1} \prod_q (1 + q^{-s} + q^{-2s})^{-1} \prod_r (1 - r^{-s})^{-2}$$

where $p \equiv$ quadratic non-residue mod 23

$q \equiv$ quadratic residue mod 23 of the form $23A^2 + B^2$

$r \equiv$ quadratic residue mod 23 not of the form $23A^2 + B^2$

(4) Let $\sum_{n=1}^{\infty} \lambda(n) \cdot x^n = x \prod_{n=1}^{\infty} (1 - x^n)^2 (1 - x^{11n})^2$. Then

$$\sum_{n=1}^{\infty} \lambda(n) \cdot n^{-s} = (1 - 11^{-s})^{-1} \prod_{p \neq 11} (1 - \lambda(p) \cdot p^{-s} + p^{1-2s})^{-1}$$

(Ramanujan says that "$\lambda(p)$ can be determined." [8]).

(5) Let $\sum_{n=1}^{\infty} \lambda(n) \cdot x^n = x f^3(-x) f^3(-x^7)$ where $f(-x) = \prod_{n=1}^{\infty} (1 - x^n)$

Then $\sum_{n=1}^{\infty} \lambda(n) \cdot n^{-s} = (1 + 7^{-s})^{-1} \prod_p (1 - p^{2-2s})^{-1}$

$$\prod_q (1 + 2c_q \cdot q^{-s} + q^{2-2s})^{-1}$$

where $q = 7u^2 + v^2$, $c_q = 7u^2 - v^2$ and $q \equiv 1, 2, 4 \pmod 7$ and $p \equiv 3, 5, 6 \pmod 7$.

(*c*) Euler product developments as listed in Birch [1]. With the usual notation for Eisenstein series, writing η instead of f as in [1], the results as given by Birch are as follows :

I. Suppose that A and B are any two integers such that $A^2 + 3B^2 = p$ and $A \equiv 1 \pmod 3$, p being a prime of the form $6k + 1$. Let

$\Sigma a_0(n)\, q^{n/6} = \eta^4$, $\Sigma a_1(n) \cdot q^{n/6} = \eta^4 P$,

$\Sigma a_2(n) \cdot q^{n/6} = \eta^4 Q$, $\Sigma a_3(n)\, q^{n/6} = \eta^4 R$,

$\Sigma a_4(n)\, q^{n/6} = \eta^4 Q^2 + 288\sqrt{70}\, \eta^{20}$, $\Sigma a_5(n)\, q^{n/6} = \eta^4 QR$

$\Sigma a_7(n)\, q^{n/6} = \eta^4 Q^2 R + 10080\sqrt{286}\, \eta^{20} R$.

In all these cases

$$\sum_{n=1}^{\infty} a_K(n) \cdot n^{-s} = \prod_p (1 - a_K(p) \cdot p^{-s} + p^{2K+1-2s})^{-1} \text{ where } p$$

assumes all prime values greater than 3.

If $K = 0, 2, 3, 5$ then $a_K(p) = 0$ for $p \equiv 5\ (6)$,

$a_K(p) = (A + \sqrt{-3} \cdot B)^{2K+1} + (A - \sqrt{-3}\, B)^{2K+1}$

for $p \equiv 1 \pmod 6$. For all values of n, $a_1(n) = n a_0(n)$. But $a_4(p)$ and $a_7(p)$ do not seem to have such simple laws.

II. Suppose again that A and B are defined as in I. and let

$$\Sigma a_0(n)\ q^{n/3} = \eta^8, \quad \Sigma a_1(n)\ q^{n/3} = \eta^8\ p$$

$$\Sigma a_2(n)\ q^{n/3} = \eta^8\ Q + 6\sqrt{10}\ \eta^{16}, \quad \Sigma a_3(n)\ q^{n/3} = \eta^8\ R$$

$$\Sigma a_4(n)\ q^{n/3} = \eta^8\ Q^2 + 6\sqrt{70}\ \eta^{16}\ Q, \quad \Sigma a_5(n)\ q^{n/3} = \eta^8\ QR + 12\sqrt{55}\ \eta^{16}\ R$$

$$\Sigma a_7(n)\ q^{n/3} = \eta^8\ Q^2 R + 12\sqrt{910}\ \eta^{16}\ QR.$$

In all these cases

$$\sum_{n=1}^{\infty} a_K(n) \cdot n^{-s} = \prod_p (1 - a_K(p) \cdot p^{-s} + p^{2K+3-2s})^{-1}$$

where p assumes all prime values except 3.

If $K = 0$ or 3, then $a_K(p) = 0$ for $p \equiv 2 \pmod 3$,

$a_K(p) = (A + \sqrt{-3} \cdot B)^{2K+3} + (A - \sqrt{-3} \cdot B)^{2K+3}$ for $p \equiv 1\ (3)$,

$a_1(n) = n\ a_0(n)$.

III. Suppose that C and D are integers such that $C^2 + 4D^2 = p$ where p is of the form $4k + 1$.

$$\Sigma a_0(n)\ q^{n/4} = \eta^6, \quad \Sigma a_1(n)\ q^{n/4} = \eta^6\ p, \quad \Sigma a_2(n)\ q^{n/4} = \eta^6\ Q,$$

$$\Sigma a_3(n)\ q^{n/4} = \eta^6\ R + 24i\sqrt{35}\ \eta^{18}, \quad \Sigma\ a_4(n)\ q^{n/4} = \eta^6\ Q^2,$$

$$\Sigma a_5(n)\ q^{n/4} = \eta^6\ QR + 24i\ \sqrt{1155}\ \eta^{18}\ Q,$$

$$\Sigma a_7(n)\ q^{n/4} = \eta^6\ Q^2 R + 120i\ \sqrt{3003}\ \eta^{18}\ Q^2 \text{ where } i^2 = -1.$$

In all these cases $\sum_{n=1}^{\infty} a_K(n) \cdot n^{-s} = \Pi_1 \Pi_2$ where

$\Pi_1 = \prod_p (1 - a_K(p) \cdot p^{-s} - p^{2K+2-2s})^{-1}$, p assuming primes of all the form $4k - 1$ and

$\Pi_2 = \prod_p (1 - a_K(p) \cdot p^{-s} + p^{2K+2-2s})^{-1}$, p assuming all primes of the form $4k + 1$. If $K = 0, 2$ or 4, then $a_K(p) = 0$ for $p \equiv 3 \pmod 4$, $a_K(p) = (C + 2iD)^{2K+2} + (C - 2iD)^{2K+2}$ for $p \equiv 1 \pmod 4$, $a_1(n) = n\ a_0(n)$.

IV. Let $\Sigma a_0(n)\ q^{n/12} = \eta^2, \ \Sigma a_1(n)\ q^{n/12} = \eta^2\ P$

$$\Sigma a_2(n)\ q^{n/12} = \eta^2\ Q + 48\ \eta^{10}, \quad \Sigma a_3(n)\ q^{n/12} = \eta^2\ R + 360\ i\sqrt{3}\ \eta^{14}$$

$$\Sigma a_4(n)\ q^{n/12} = \eta^2\ Q^2 + 672\ \eta^{10}\ Q,$$

$$\Sigma a_5(n)\ q^{n/12} = \eta^2\ QR + 96\sqrt{1045}\ \eta^{10}\ R + 216\ i\sqrt{7315}\ \eta^{14}\ Q + 103680\ i\ \sqrt{7}\ \eta^{22}.$$

$$\Sigma a_7(n)\ q^{n/12} = \eta^2\ Q^2 R + 48\sqrt{910 \cdot 2911}\ \eta^{10}\ QR + 216\ i\sqrt{5005 \cdot 2911}\ \eta^{14}\ Q^2 - 471744\ i\sqrt{22}\ \eta^{22}\ Q.$$

There is a minor difference between Ramanujan's original manuscript and the version appearing in [1] as given above. All the quadratic (real and imaginary) irrationalities and the rational coefficients ($\neq 1$) appear with signs ± 1 in Ramanujan's version.

The following is an additional list of Euler product developments found in [8] which is incomplete except for the first one which is already quoted above.

(1) If $\quad \Sigma\, a(n)\, x^n = x^2 \Pi (1 - x^{3n})^{16}$ and

$$\Sigma\, A(n)\, x^n = x\, \Pi (1 - x^{3n})^8 \cdot \left(1 + 240 \sum_1^\infty \frac{n^3 x^{3n}}{1 - x^{3n}}\right)$$

$$\Sigma \frac{A(n) + 6a_n\sqrt{10}}{n^s} = (1 - 6\sqrt{10} \cdot 2^{-s} + 2^{7-2s})^{-1} \;\; (1 + 96\sqrt{10} \cdot 5^{-s}$$
$$+ 5^{7-2s})^{-1} (1 - 260 \cdot 7^{-s} + 7^{7-2s})^{-1}$$
$$\times (1 + 1920\sqrt{10} \cdot 11^{-s} + 11^{7-2s})^{-1}$$

(this is the same as $\Sigma a_2(n)\, n^{-s}$ in II).

(2) If $\quad \Sigma\, a(n)\, x^n = x^3 \Pi (1 - x^{4n})^{18}$. Then

$$156\, \Sigma\, a(n) \cdot n^{-s} = (1 - 78 \cdot 3^{-s} + 3^{8-2s})^{-1} \;\; (1 + 510 \cdot 5^{-s} + 5^{8-2s})^{-1}.$$
$$(1 + 1404 \cdot 7^{-s} + 7^{8-2s})^{-1} \cdots - (1 + 78 \cdot 3^{-s} + 3^{8-2s})^{-1} (1 + 510 \cdot 5^{-s}$$
$$+ 5^{8-2s})^{-1} (1 + 1404 \cdot 7^{-s} + 7^{8-2s})^{-1} \ldots$$

"Presumably there are analogous results for

$$x^5 \Pi (1 - x^{12n})^{10},\; x^7 \Pi (1 - x^{12n})^{14},\; x^5 \Pi (1 - x^{6n})^{20} \text{ and } x^{11} \Pi (1 - x^{12n})^{22}."$$

As stated in the introduction, we shall prove the results in (*c*) and later mention about the results in (*a*) and (*b*) and their connection with the work of Hecke and others. Since we shall be using the results of Rankin in [9], we shall summarize them briefly.

3. Rankin's results

Let $\Gamma(1)$ be the full modular group $SL(2, ZZ)$ and $\Gamma'(1)$ its commutator subgroup. Then $\Gamma(1)/\Gamma'(1)$ is cyclic of order 12 and $\Gamma'(1) \supset \Gamma(12)$ (the principal congruence subgroup of stufe 12). For each real dimension $-k$, there exist 6 possible "multiplier systems" on $\Gamma(1)$ which we denote by $v^{(r)}$, $r \in \mathcal{R} = \{0, 4, 6, 8, 10, 14\}$ defined as follows. If

$$U = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \text{ and } V = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \text{then} \quad v^{(r)}(U) = \exp\, [\pi i (k - r)]/6,$$
$$v^{(r)}(V) = \exp\, [-\pi i (k - r)]/2.$$

We observe that the six multiplier systems agree on $\Gamma'(1)$ and we take $v^{(0)} = u$ and set $M_k = \{\Gamma'(1), k, u\}_0 =$ the space of cusp forsm for $\Gamma'(1)$ of weight k and multiplier system u, and $M_k^{(r)} = \{\Gamma(1), k, v^{(r)}\}_0 =$ the space of cusp forms for $\Gamma(1)$ of weight k and multiplier system $v^{(r)}$,

Then $M^{(r)}{}_k \neq 0$ only if $k \geqslant r$ and a basis for $M^{(r)}{}_k$ is given by $E_{r+12s}\,\eta^{2(k-r)-12s}$ $(0 \leqslant s \leqslant k - r/12)$ except when $k \equiv r \pmod{12}$ and $s = (k-r)/12$. Here E_s denotes the usual Eisenstein series $\epsilon\{\Gamma(1), v, 1\}$. (Later we shall use Ramanujan's notation Q and R for E_4 and E_6).

Now, $M_k = \bigoplus_{r \in \mathscr{R}} M_k{}^{(r)}$ (direct sum).

$M_k \subset \{\Gamma(12), k, \epsilon\}_0$: the space of cusp forms for the congruence subgroup $\Gamma(12)$ of weight k and character ϵ defined by

$$\epsilon(n) = \left(\sin \frac{\pi n}{2}\right)^k.$$

$$\text{dim. } M_k = \delta_k = \max\left(1, -1 - \left[-\frac{k}{2}\right]\right).$$

For each positive divisor t of 12, we denote by $M_k^{(r)}(t)$, the space of cusp forms in whose Fourier expansion only powers $\exp(2\pi imz/12)$ with $(m, 12) = t$ occur. In other words, if $f \in M_k^{(r)}(t)$ and $q = \exp(2\pi iz)$, $f = \sum_{(n,12)=1} a_n (q^n)^{12}$. They are of divisor t in the sense of Hecke. We notice that $(k-r, 12) = t$ in case $f \in M_k^{(r)}(t)$ and conversely if $(k-r, 12) = t$ then $f \in M_k^{(r)}(t)$. It follows from Hecke theory that $M_k^{(r)}(t)$ are stable under Hecke operators T_n with $(n, 12) = 1$. We also observe that if $F \in M_k^{(r)}$, $F(z) = \sum_{n=1}^{\infty} \lambda(n)\, q^{nt/12}$ with $\lambda(1) = 1$, is a normalized eigenfunction for the Hecke operators T_n, then the Dirichlet series $\sum_{n=1}^{\infty} \lambda(n) \cdot n^{-s}$ admits an Euler product of the form $\prod_{p \times t} (1 - \lambda(p) \cdot p^{-s} + \epsilon(p) \cdot p^{k-1-2s})^{-1}$ and conversely. A necessary condition for F to be an eigenfunction of $T_s(p \times t)$ is given by

$$\lambda(p)^2 = \lambda(p^2) + p^{k-1} \cdot \epsilon(p) \quad (\text{since } \lambda(1) = 1). \tag{1}$$

For the proof of Ramanujan's results in [1] this will be useful.

From the discussion preceding the statements of Ramanujan's results, we see that all Euler products of Ramanujan correspond to cusp forms in $M_k(t)$ for different divisors t of 12 ($t = 2$ in Case I, $t = 4$ in Case II, $t = 3$ in Case III, $t = 1$ in Case IV). The case of $t = 12$ occurs in his results in [7] although not in great detail. The case $t = 6$ occurs only in one example of Ramanujan. [8].

We shall therefore tabulate below the different values of t and corresponding values of k and the type of congruence subgroups and character to which the forms belong. This is followed by tables of bases of $M_k(t)$ $(t \mid 12)$ and $k \leqslant 30$ in case $t = 12$ and $k \leqslant 20$ in other cases, since these are the only cases relevant to Ramanujan's results. In the following $F(z) \in M_k(t)$, $\epsilon(n) = (\sin \pi n/2)^k$.

4. Proof of Ramanujan's results

In this paragraph, we shall prove all statements of § 2, regarding the existence of Euler products. In § 5, we shall take up the results of § 2 concerning the structure of Euler products, since, they overlap with some of the results in [7].

We shall consider the different cases as set out in table 1 (see table 2).

Table 1.

Cases	t	Fourier expansion of $F(z)$	Condition on k	Type
1.	12	$F(z) = \Sigma a_n q^n$	k even, $\neq 14$, $\geqslant 12$	$\{\Gamma(1), k, 1\}_0$
2.	6	$F(z) = \Sigma a_n q^{n/2}$	k even, $\neq 8$, $\geqslant 6$	$\{\Gamma(2), k, 1\}_0$
3.	4	$F(z) = \Sigma a_n q^{n/3}$	k even, $\neq 6$, $\geqslant 4$	$\{\Gamma(3), k, 1\}_0$
4.	3	$F(z) = \Sigma a_n q^{n/4}$	k odd, $\neq 5$, $\geqslant 3$	$\{\Gamma(4), k, \epsilon\}_0$
5.	2	$F(z) = \Sigma a_n q^{n/6}$	k even, $\neq 4$, $\geqslant 2$	$\{\Gamma(6), k, 1\}_0$
6.	1	$F(z) = \Sigma a_n q^{n/12}$	k odd, $\neq 3$, $\geqslant 1$	$\{\Gamma(12), k, \epsilon\}_0$

Table 2.

Cases	t	k	$\{r \in \mathcal{R}\ (k-r, 12) = t\}$	Basis of $M_k(t)$
1.	12	12, 16, 18, 20, 22, 24, 26, 28, 30	$\{0\}, \{4\}, \{6\}, \{8\}, \{10\}, \{0\}, \{14\}, \{4\}, \{6\}$	$\{\Delta\}, \{\Delta Q\}, \{\Delta R\}, \{\Delta Q^2\}, \{\Delta QR\}, \{\Delta E_{12}, \Delta^2\}, \{\Delta Q^2 R\}, \{\Delta^2 Q, \Delta E_{16}\}, \{\Delta^2 R, \Delta E_{18}\}$
2.	6	6, 10, 12, 14, 16, 18, 20	$\{0\}, \{4\}, \{6\}, \{8\}, \{10\}, \{0\}, \{14\}$	$\{\eta^{12}\}, \{\eta^{12} Q\}, \{\eta^{12} R\}, \{\eta^{12} Q^2\}, \{\eta^{12} QR\}, \{\eta^{12} E_{12}, \eta^{36}\}, (\eta^{12} Q^2 R\}$
3.	4	4, 8, 10, 12, 14, 16, 18	$\{0\}, \{0, 4\}, \{6\}, \{4, 8\}$ $\{6\ 10\}, \{0\ 8\}, \{10, 14\}$	$\{\eta^8\}, \{\eta^{16}, \eta^8 Q\}, \{\eta^8 R\}, \{\eta^{16} Q, \eta^8 Q^2\}, \{\eta^{16} R, \eta^8 QR\}, \{\eta^{32}, \eta^{16} Q^2, \eta^8 E_{12}\}, \{\eta^{16} QR, \eta^8 Q^2 R\}$
4.	3	3, 7, 9, 11, 13, 15, 17	$\{0\}, \{4\}, \{0, 6\}, \{8\}$ $\{4, 10\}, \{0, 6\}, \{8, 14\}$	$\{\eta^6\}, \{\eta^6 Q\}, \{\eta^{18}, \eta^6 R\}, \{\eta^6 Q^2\}, \{\eta^{18} Q, \eta^6 QR\}, \{\eta^{30}, \eta^{18} R, \eta^6 E_{12}\}, \{\eta^{18} Q^2, \eta^6 Q^2 R\}$
5.	2	2, 6. 8, 10, 12, 14, 16	$\{0\}, \{4\}. \{6\}, \{0, 8\}, \{10\}, \{0, 4\}\ \{6, 14\}$	$\{\eta^4\}, \{\eta^4 Q\}, \{\eta^4 R\}, \{\eta^4 Q^2, \eta^{20}\}, \{\eta^4 QR\}, \{\eta^4 E_{12}, \eta^{20} Q, \eta^{28}\}, \{\eta^4 Q^2 R, \eta^{20} R\}$.
6.	1	1, 5, 7, 9, 11, 13, 15	$\{0\}, \{0, 4\}, \{0, 6\}, \{4, 8\}, \{0, 4, 6, 10\}, \{0, 6, 8\}, \{4, 8, 10, 14\}$	$\{\eta^2\}, \{\eta^{10}, \eta^2 Q\}, \{\eta^{14}, \eta^2 R\}, \{\eta^{10} Q, \eta^2 Q^2\}, \{\eta^{22}, \eta^{14} Q, \eta^{10} R, \eta^2 QR\}, \{\eta^2 \Delta, \eta^2 E_{12}, \eta^{14} R, \eta^{10} Q^2\}, \{\eta^{22} Q, \eta^{14} Q^2, \eta^{10} QR, \eta^2 Q^2 R\}$

Case 1: $t = 12$, $F(z) = \sum_{n=1}^{\infty} a_n q^n \in \{\Gamma(1), k, 1\}_0$.

Except in the case of weights 24, 28, 30, the space $M_k(12)$ is 1-dimensional and hence the basis forms are normalized eigenfunctions (since $\lambda(1) = 1$) in all these cases. They all occur in [7]. The case of $k = 24$ does not occur explicitly in Ramanujan, except in his reference to it in his letter to Hardy (Appendix [6]). In this case we see easily that $\triangle E_{12} + \lambda\triangle^2$ is an eigenfunction and the constant λ can be determined. In fact Hecke says [2] $\lambda \in Q(\sqrt{144169})$ as can be verified. λ satisfies the quadratic equation $\lambda^2 + \lambda(2\mu - 1128) + (\mu^2 - 128272\mu + 2048 - 2^{23}) = 0$ (where $\mu = 65520/691$). The discriminant of this equation is $24^2 \times 144169$. In the case of weights 28 and 30, Ramanujan says [8] $\triangle^2 Q$ and $\triangle^2 R$ do not have Euler products, "however they are differences of two such products". This is obvious from the fact that $\triangle E_{16} + \alpha_1\triangle^2 Q$ and $\triangle E_{16} + \alpha_2\triangle^2 Q$ are Euler products and similarly $\triangle E_{18} + \beta_1\triangle^2 R$ and $\triangle E_{18} + \beta_2\triangle^2 R$ are Euler products for some constants $\alpha_1, \alpha_2, \beta_1, \beta_2$ which can be determined.

Case 2. $t = 6$. $F(z) = \Sigma\, a_n\, q^{n/2} \in \{\Gamma(2), k, 1\}_0$.

As in case (1), for all k except 18, the eigenfunctions are given by $\eta^{12}\, E_\nu$ where E_ν is an Eisenstein series of suitable weight. Only in the case of $k = 18$, the eigenfunctions are of the form $\eta^{12} E_{12} \pm \alpha\eta^{36}$ where α can be determined using (1). The only results of Ramanujan in this case are for the weight $k = 6$ in [7] and η^{36} (in the case $k = 18$) in his indirect reference in the letter [6].

Case 3. $t = 4$. $F(z) = \Sigma\, a_n\, q^{n/3} \in \{\Gamma(3), k, 1\}_0$.

It is easily seen that the above case refers to II of Ramanujan's results stated before. We observe first the Euler product for η^8 was already stated in [7] and proved by Mordell [4]. The function $\eta^8 P$ is not a cusp form, however using the operator $q\, d/dq$, one sees that $a_1(n) = n\, a_0(n)$ as in [7], and this yields the formal Euler product development for $\eta^8 P$ from that of η^8. In the case of $k = 10$, since dim $M_{10}(4) = 1$ and $\eta^8 R$ is normalized, the Dirichlet series of $\eta^8 R$ has an Euler product. In the case of weights $k \geqslant 8$ ($k \neq 10$) any eigenfunction is a linear combination of the forms given below and we need to evaluate the constants and compare them with the values given by Ramanujan. This is done by using the equation (1) for special primes p. The Fourier coefficients $a_0(n)$ can be taken from the table of Newman [5] (in fact $a_0(n) = 0$ if $4 \times n$ and $a_0(4m) = p_8(m)$ in Newman's notation). For the computation of Fourier coefficients of η^α, we also use Newman's tables [5]. By explicit computation we can evaluate $a_K(2)$ and $a_K(4)$ for various K. The identity we need is the following:

$$(a_K(2))^2 = a_K(4) + a_K(1)\cdot 2^{k-1}$$

$$= a_K(4) + 2^{2K+3} \text{ since } a_K(1) = 1 \text{ and } k = 2K + 4.$$

(1) $\underline{k = 8,\ K = 2.}$ $F(z) = \Sigma\, a_2(n)\, q^{n/3} = \eta^8 Q + \lambda_2\eta^{16}$ is a normalized eigenfunction for some value of λ_2 since $a_2(1) = 1$ and the term $q^{11/3}$ occurs only in $\eta^8 Q$. Then $a_2(2) = \lambda_2$ and

$$\lambda_2^2 = (a_2(2))^2 = a_2(4) + 2^7 = 360, \text{ i.e., } \lambda_2 = \pm 6\sqrt{10}.$$

(2) $\underline{k=12}$, $\underline{K=4}$, $F(z) = \Sigma\, a_4(n)\, q^{n/3}$

$$= \eta^8 Q^2 + \lambda_4 \eta^{16} Q \text{ is an eigenfunction.}$$

$$a_4(2)^2 = \lambda_4^2 = a_4(4) + 2^{11}.$$

$$= 472 + 2^{11} = 2520,\ \lambda_4 = \pm 6\sqrt{70}.$$

(3) $\underline{k=14}$, $\underline{K=5}$, $F(z) = \Sigma a_5(n)\, q^{n/3}$

$$= \eta^8 QR + \lambda_5 \eta^{16} R \text{ is an eigenfunction}$$

$$a_5(2)^2 = \lambda_5^2 = a_5(4) + 2^{13}$$

$$= -272 + 2^{13} = 7920,\quad \lambda_5 = \pm 12\sqrt{55}.$$

(4) $\underline{k=18}$, $\underline{K=7}$, $F(z) = \Sigma\, a_7(n)\, q^{n/3}$

$$= \eta^8 Q^2 R + \lambda_7 \eta^{16} QR \text{ is an eigenfunction}$$

$$a_7(2)^2 = \lambda_7^2 = a_7(4) + 2^{17}$$

$$= -32 + 2^{17} = 144 \times 910,\ \lambda_7 = \pm 12\sqrt{910}.$$

This exhausts Ramanujan's list. As can be seen, Ramanujan does not consider the case $K=6$ corresponding to $k=16$ although a reference to η^{32} is made in his letter to Hardy [6]. The eigenfunction is of the form $\eta^{32} + \lambda_6 \eta^{16} Q^2 + \mu_6 \eta^8 E_{12}$ for some constants λ_6 and μ_6. They can be evaluated by the above method.

Case 4. $\underline{t=3.}$ $F(z) = \Sigma\, a_n q^{n/4} \in \{\Gamma(4),\ k,\ \epsilon\}_0$.

This case refers to Ramanujan's results in III, stated above. As in II, the Euler product for η^6 was already stated by Ramanujan in [7] and proved by Mordell [4]. The function $\eta^6 P$ is not a cusp form, however the relation $a_1(n) = n\, a_0(n)$ and the formal Euler product for $\eta^6 P$ can be derived as before. In the case of weights 7 and 11, the space is 1-dimensional and since $\eta^6 Q$ and $\eta^6 Q^2$ are normalized, the corresponding Dirichlet series have Euler products. For weights $k = 9$, 13, 15, 17, any eigenfunction is a linear combination of the basis forms as given below and we need to evaluate the constants and compare them with the values given by Ramanujan. This is done by using the identity (1) for the prime $p=3$. Observe that $a_0(n) = 0$ if n is even or odd and $\neq 1 \pmod 4$. In fact, if $n = 4m+1$, $a_0(n) = p_6(m)$ in Newman's notation. As before we explicitly evaluate $a_K(3)$ and $a_K(9)$ for various K.

The identity we need is the following:

$$a_K(3)^2 = a_K(9) + \epsilon(3)\cdot 4^{k-1}\cdot a_K(1)$$

$$= a_K(9) - 2^{2K+2} \text{ since } a_K(1) = 1,\ \epsilon(3) = -1 \text{ and } k = 2K+3.$$

(1) $\underline{k=9,\ K=3.}$ $F(z) = \Sigma a_3(n)\, q^{n/4}$

$$= \eta^6 R + \lambda_3 \eta^{18} \text{ is a normalized eigenfunction for some}$$

value of λ_3, since $a_3(1) = 1$ and the term $q^{1/4}$ occurs only in $\eta^6 R$.

$$a_3(3)^2 = \lambda_3^2 = a_3(9) - 3^8$$

$$= -24^2\cdot 35,\quad \lambda_3 = \pm 24i\sqrt{35}.$$

(2) $k = 13,\ K = 5.$ $F(z) = \Sigma\, a_5(n)\, q^{n/4}$

$$= \eta^6 QR + \lambda_5 \eta^{18} Q \text{ is an eigenfunction.}$$

$$a_5(3)^2 = \lambda_5^2 = a_5(9) - 3^{12}$$
$$= -24^2 \cdot 1155,\quad \lambda_5 = \pm 24i\sqrt{1155}$$

(3) $k = 17,\ K = 7.$ $F(z) = \Sigma\, a_7(n)\, q^{n/4}$

$$= \eta^6 Q^2 R + \lambda_7 \eta^{18} Q^2 \text{ is an eigenfunction.}$$

$$a_7(3)^2 = \lambda_7^2 = a_7(9) - 3^{16}$$
$$= -120^2 \cdot 3003,\ \lambda_7 = \pm 120i\sqrt{3003}.$$

As before, we observe that the case $K = 6$ (corresponding to $k = 13$ is not found in Ramanujan's results although there is an indirect reference to η^{30} in his letter. In this case the eigenfunction is of the form $\eta^6 E_{12} + \lambda_6 \eta^{18} R + \mu_6 \eta^{30}$ and the constants λ_6 and μ_6 can be evaluated as above.

Case 5 : $t = 2,\ F(z) = \Sigma\, a_n\, q^{n/6} \in \{\Gamma(6),\ k,\ 1\}_0$

This case refers to Ramanujan's results in I, stated above. As in other cases, the Euler product for η^4 was already stated by Ramanujan in [7] and proved by Mordell. The function $\eta^4 P$ is not a cusp form and as before we can derive $a_1(n) = na_0(n)$ and the formal Euler product for $\eta^4 P$ from that of η^4. In the case of weights 6, 8, 12 the space $M_k(2)$ is 1-dimensional and since $\eta^4 Q$, $\eta^4 R$, $\eta^4 QR$ are normalized, the corresponding Dirichlet series have Euler products. For $k = 10$, 14, 16 any eigenfunction is a linear combination of the basis forms as given below and we evaluate the constants and compare them with the values given by Ramanujan. This is done by using (1) for the prime $p = 5$. Observe that $a_0(n) = 0$ unless $n \equiv 1 \pmod 6$ and if $n = 6m + 1$, $a_0(n) = p_4(m)$ in Newman's notation. We shall explicitly evaluate $a_K(5)$ and $a_K(25)$ for different K.

The identity we need is the following:

$$a_K(5)^2 = a_K(25) + 5^{k-1} \cdot a_K(1)$$
$$= a_K(25) + 5^{2K+1} \text{ since } a_K(1) = 1 \text{ and } k = 2K + 2.$$

(1) $k = 10,\ K = 4.$ $F(z) = \Sigma\, a_4(n)\, q^{n/6}$

$= \eta^4 Q^2 + \lambda_4 \eta^{20}$ is a normalized eigenfunction for some value of λ_4, since $a_4(1) = 1$ and the term $q^{1/6}$ occurs only in $\eta^4 Q^2$.

$$\lambda_4^2 = a_4(5)^2 = a_4(25) + 5^9.$$
$$= 288^2 \cdot 70,\quad \lambda_4 = \pm 288\sqrt{70}.$$

(2) $k = 16,\ K = 7.$ $F(z) = \Sigma\, a_7(n)\, q^{n/6}$

$= \eta^4 Q^2 R + \lambda_7 \eta^{20} R$ is a normalized eigenfunction for some value of λ_7 since $a_7(1) = 1$ and the term $q^{1/6}$ occurs only in $\eta^4 Q^2 R$.

$$a_7(5)^2 = \lambda_7^2 = a_7(25) + 5^{15}$$
$$= (10080)^2 \cdot 286, \text{ i.e., } \lambda_7 = \pm 10080\sqrt{286}.$$

As before, the case $K = 6$ (corresponding to $k = 14$) is missing from Ramanujan's results although an indirect reference to η^{28} occurs in his letter [6]. From our table, we can see that the eigenfunction is of the form $\eta^4 E_{12} + \lambda_6 \eta^{20} Q + \mu_6 \eta^{28}$ and the constants λ_6 and μ_6 can be determined as before.

Case 6. $t = 1$. $F(z) = \Sigma a_n q^{n/12} \in \{\Gamma(12), k, \epsilon\}_0$.

This case refers to Ramanujan's results in IV quoted above. Although no explicit mention is made about the Euler products, in the context it is obvious that Ramanujan states that $\sum_{n=1}^{\infty} a_K(n) \cdot n^{-s}$ are all Euler products. As before, the Euler product for η^2 already occurs in [7] and proved by Mordell [4]. The function $\eta^2 P$ is not a cusp form; however, from the equation $a_1(n) = na_0(n)$ (derived as before), the formal Euler product for $\eta^2 P$ follows from that of η^2. In all the other cases, the eigenfunction is a linear combination of the basis forms as given below and we need to evaluate the constants and compare them with the values given by Ramanujan. This is done by using (1) for the primes $p = 5, 7$ and 11 in the different cases.

[Observe that $a_0(n) = 0$ if n is even or odd with $n \not\equiv 1 \pmod{12}$. In fact $a_0(n) = p_2(m)$ if $n = 12m + 1$ in Newman's notation]. The constants $a_K(n)$ are explicitly evaluated for various values of K and n.

(1) $\underline{k = 5,\ K = 2}$. $F(z) = \Sigma a_2(n) q^{n/12}$

$= \eta^2 Q + \lambda_2 \eta^{10}$ is a normalized eigenfunction for some value of λ_2, since $a_2(1) = 1$ and the term $q^{1/12}$ occurs only in $\eta^2 Q$.

$$a_2(5)^2 = \lambda_2^2 = a_2(25) + 5^4 \quad (\text{since } \epsilon(5) = 1).$$
$$= 1679 + 5^4 = 2304,\ \lambda_2 = \pm 48.$$

(2) $\underline{k = 7,\ K = 3}$. $F(z) = \Sigma a_3(n) q^{n/12}$

$= \eta^2 R + \lambda_3 \eta^{14}$ is a normalized eigenfunction for some value of λ_3, since $a_3(1) = 1$ and the term $q^{1/12}$ occurs only in $\eta^2 R$.

$$a_3(7)^2 = \lambda_3^2 = a_3(49) - 7^6 \quad (\text{since } \epsilon(7) = -1)$$
$$= -3 \cdot (360)^2 \text{ i.e., } \lambda_3 = \pm 360i\sqrt{3}.$$

(3) $\underline{k = 9,\ K = 4}$. $F(z) = \Sigma a_4(n) q^{n/12} = \eta^2 Q^2 + \lambda_4 \eta^{10} Q$

$$a_4(5) = \lambda_4 \quad a_4(25) = 60959$$
$$\lambda_4^2 = a_4(25) + 5^8 = 451584, \text{ i.e., } \lambda_4 = \pm 672.$$

(4) $\underline{k = 11,\ K = 5}$. $F(z) = \Sigma a_5(n) q^{n/12}$

$$= \eta^2 QR + \lambda_5 \eta^{10} R + \mu_5 \eta^{14} Q + \nu_5 \eta^{22}$$

is a normalized eigenfunction for some values λ_5, μ_5, ν_5, since $a_5(1) = 1$ and the term $q^{1/12}$ occurs only in $\eta^2 QR$.

(a) $\lambda_5^2 = a_5(5)^2 = a_5(25) + 5^{10} = 96^2 \times 1045$, i.e., $\lambda_5 = \pm 96\sqrt{1045}$

(b) $\mu_5^2 = a_5(7)^2 = a_5(49) - 7^{10} = -216^2 \cdot 7315$, i.e., $\mu_5 = \pm 216i\sqrt{7315}$

(c) $\nu_5^2 = a_5(11)^2 = a_5(121) - 11^{10} = -7 \cdot (103680)^2$, i.e., $\nu_5 = \pm 103680i\sqrt{7}$.

(4) $k = 15$, $K = 7$. $F(z) = \Sigma\, a_7(n)\, q^{n/12}$

$$= \eta^2 Q^2 R + \lambda_7\, \eta^{10}\, QR + \mu_7\, \eta^{14}\, Q^2 + \nu_7\, \eta^{22}\, Q$$

is a normalized eigenfunction for some values λ_7, μ_7, ν_7, since $a_7(1) = 1$ and the term $q^{1/12}$ occurs only in $\eta^2\, Q^2 R$.

(a) $\lambda_7^2 = a_7(5)^2 = a_7(25) + 5^{14} = 48^2 \times 910 \times 2911$;
$\lambda_7 = \pm\, 48\, \sqrt{910 \cdot 2911}$

(b) $\mu_7^2 = a_7(7)^2 = a_7(49) - 7^{14} = -\, 216^2 \times 5005 \times 2911$;
$\mu_7 = \pm\, 216 i \sqrt{5005 \cdot 291}$

(c) $\nu_7^2 = a_7(11)^2 = a_7(121) - 11^{14} = -\, 2200 \times (471744)^2$;
$\nu_7 = \pm\, 4717440 i\, \sqrt{22}.$

The computations in (3) and (4) were carried out in the TIFR Computer. Notice that all the values for the constants agree with those given by Ramanujan except ν_7, where the difference is a factor of 10. Again as in other cases, $K = 6$ (corresponding to $K = 13$) is missing from Ramanujan's results. We see from our table that the eigenfunction is of the form $\eta^2 E_{12} + \lambda_6 \eta^2 \Delta + \mu_6 \eta^{10}\, Q^2 + \nu_6\, \eta^{14} R$ for some constants λ_6, μ_6, ν_6 which can be evaluated as above.

5. Explicit form of Euler products of § 2c

In this paragraph, we shall prove the results of § 2c on the explicit form of Euler products.

All the Euler products mentioned in § 2c are Hecke L-series attached to $Q(\sqrt{-3})$ and $Q(\sqrt{-1})$. In fact we shall see that the Euler products in Cases I and II of Ramanujan (whenever the dimension is 1) are Hecke L-series attached to $Q(\sqrt{-3})$ and those in Case III (whenever the dimension is 1) are Hecke L-series attached to $Q(\sqrt{-1})$. These Dirichlet series are in turn Mellin transforms of theta series $\theta_k(\tau, \rho, \theta, Q\sqrt{D})$ introduced by Hecke (23, [2]). We briefly recall their definition and properties.

Let $\varphi(\sqrt{D})$ be an imaginary quadratic field with discriminant D, θ, an integral ideal in $\varphi(\sqrt{D})$, $\rho \in \theta$ and Q, a positive rational integer. Let $A = |\, N\mathfrak{a}\,|$. Then for every natural number k, we form the following binary theta series:

$$\theta_k(\tau, \rho, \mathfrak{a}, Q\sqrt{D}) = \sum_{\mu \equiv \rho\ (\mathrm{mod}\ \mathfrak{a}\, Q\sqrt{D})} \mu^{k-1} \exp\left(2\pi i\tau \frac{N\mu}{AQ\,|\,D\,|}\right)$$

Since $\mathfrak{a}$ and Q are fixed throughout, we may denote this for simplicity, $\theta_k(\tau, \rho)$. We have then the following relations.

(1) $\theta_k(\tau, \rho\epsilon) = \epsilon^{k-1} \cdot \theta_k(\tau, \rho)$ for every unit ϵ in $Q(\sqrt{D})$

(2) $\theta_k(\tau + 1, \rho) = \exp\,(2\pi i\,(N\rho/AQ\,|\,D\,|)\ \theta_k(\tau, \rho)$

(3) $\theta_k(-1/\tau, \rho) = \tau^k / Q\sqrt{D} \sum_{\substack{\alpha\ \mathrm{mod}\ \theta\ Q\sqrt{D} \\ \alpha \in \mathfrak{a}}} \exp\,(2\pi i\, tr\,(\alpha\bar{\rho}/AQD)\ \theta_k(\tau, \alpha).$

We shall be concerned only with $D = -3$ and $D = -4$ in this section, and special ρ and $\mathfrak{a}$.

(a) $D = -3$. (i) $\rho = 1$, $\mathfrak{a} = 1$, $Q = 1$

(ii) $\rho = 1$, $\mathfrak{a} = 1$, $Q = 2$.

(i) From properties (1)-(3) listed, we see that in this case $\theta_k(\tau, 1) \in M_k(4$ (case (3) or § 4 or II of results in § 2). The values of k for whic dim $M_k(4) = 1$, are $k = 4$ and 10. Hence the cusp forms η^8 and $\eta^8 R$ have for Mellin transforms the Hecke L-series $L(s, \chi_4, Q(\sqrt{-3}))$ and $L(s, \chi_{10}, Q(\sqrt{-3}))$ where $\chi_4((\mu)) = \mu^3$ for $\mu \equiv 1 \pmod{\sqrt{-3}}$ and $\chi_{10}((\mu)) = \mu^9$ for $\mu \equiv 1 \pmod{\sqrt{-3}}$, in agreement with Ramanujan' results in II of § 2.

(ii) In this case we may verify that $\theta_k(\tau, 1, 1, 2\sqrt{-3}) \in M_k(2)$ (case 5 of § 4 or I of results in § 2). The values of k for which dim $M_k(2) = 1$ are $k = 2, 6, 8, 12$ and the cusp forms are given by η^4, $\eta^4 Q$, $\eta^4 R$ and $\eta^4 QR$. Their Mellin transforms are given by $L(s, \chi_k, Q(\sqrt{-3})$ where $\chi_k((\mu)) = \mu^{k-1}$ for $\mu \equiv 1 \pmod{2\sqrt{-3}}$ $(k = 2, 6, 8, 12)$ in agreement with Ramanujan's results in I of § 2.

(b) $D = -4$. (i) $\rho = 1$, $\mathfrak{a} = 1$, $Q = 1$ (ii) $\rho = 2$, $\mathfrak{a} = 1$, $Q = 2$.

(i) As in the case of $D = -3$, we deduce that $\theta_k \in M_k(3)$ (Case (4) of § 4 or III of results in § 2). The values of k for which dim $M_k(3) = 1$ are $k = 3, 7, 11$ and the cusp forms are given by η^6, $\eta^6 Q$, $\eta^6 Q^2$. Their Mellin transforms are given by $L(s, \chi_k, Q(\sqrt{-1}))$ where $\chi_k((\mu)) = \mu^{k-1}$ for $\mu \equiv 1 \pmod{\sqrt{-4}}$ and $k = 3, 7, 11$. This is in agreement with Ramanujan's results in III of § 2.

(ii) In this case $\theta_k(\tau, 2, 1, 2\sqrt{-4}) \in M_k(6)$ (Case (2) of § 4) and the values of k for which dim. $M_k(6) = 1$ are $k = 6, 10, 12, 14, 16, 20$ and the cusp forms are given by η^{12}, $\eta^{12} Q$, $\eta^{12} R$, $\eta^{12} Q^2$, $\eta^{12} QR$ and, $\eta^{12} Q^2 R$. Their Mellin transforms are given by $L(s, \chi_k, Q(\sqrt{-4}))$ where $\chi_k((\mu)) = \mu^{k-1}$ for $\mu \equiv 2 \pmod{2\sqrt{-4}}$) and $k = 6, 10, 12, 14, 16, 20$. Although these Euler products do not figure in Ramanujan's results quoted in § 2, the Euler product for η^{12} may be identified with that for $\sum_{n=1}^{\infty} e_{12}(n) \,.\, n^{-s}$ (notation as in [7]) as shown by Rankin [9]. The other Euler products $\eta^{12} Q$, $\eta^{12} R$ also occur in Rankin [9].

The Euler products of Case (1) of § 4, corresponding to the full modular group, do not belong to the above category. It has been shown by Hecke that $\Delta(\tau)$ is in fact a theta-series (with spherical harmonics) attached to a quadratic form in 8 variables. (41, [2]).

The Euler product for η^2 (Case (6) of § 4 or IV of § 2) is the Mellin transform of a theta series attached to a real quadratic field. We shall take it up in the next paragraph.

6. Euler products of (a) and (b) of § 2

Regarding (a), except for $\alpha = 1, 2, 3$, all the other cases are included in § 5. The Mellin transforms of $\eta(z)$ and $\eta^3(8z)$ are given by Dirichlet L-series $L(s, \chi)$ where $\chi(n) = (3/n)$ and $(-1/n)$ respectively [4]. The Mellin transform of $\eta^2(12z)$ is the Dirichlet series $L(s, \chi, Q(4\sqrt{12}, i)/Q(i)$ which is also the Mellin transform of $\theta_+(\tau, 1, 1, \sqrt{12})$ (23, [2]).

Proof of (1)-(5). We shall take up (3) first since it is already found in the literature. [10, 11].

(3) $\eta(z)\,\eta(23z)$ is a cusp form (new) for $\Gamma_0(23)$ of nebentype belonging to the character $\epsilon(n) = (23/n)$ and weight -1. From the theorem of Serre–Deligne [11] its Mellin transfoim is an L-series $L(s, \rho)$ attached to a 2-dimensional dihedral representation ρ of $G(\bar{Q}/Q)$. More explicitly, ρ is the 2-dimensional representation induced by the non-trivial cubic character of the Galois group of the Hilbert class field $Q(\sqrt{-23}, \alpha)$ (α satisfies $x^3 - x - 1 = 0$) of $Q(\sqrt{-23})$. It can be verified by using the explicit decomposition of prime ideals of $Q(\sqrt{-23})$ in $Q(\sqrt{-23}, \alpha)$ and the definition of $L(s, \rho)$ that Ramanujan's statement (3) means precisely this.

(2) In this case $\eta(z)\,\eta(11z)$ is a cusp form of weight -1 belonging to $\Gamma_0(44)$ and character $(44/n)$ and as before its Mellin transform is $L(s, \rho)$ where ρ is the 2-dimensional dihedral representation induced by the non-trivial cubic character of Galois group of the Hilbert class field $Q(\sqrt{-11}, \beta)$ (β satisfies $x^3 - 2x^2 + 2x - 2 = 0$) of $Q(\sqrt{-11})$. It can be verified as above that Ramanujan's statement (2) is precisely this.

(1) This case is similar to (2) and (3) but slightly different. The form $\eta(z)\,\eta(7z)$ is a cusp form of weight -1 belonging to $\Gamma_0(63)$ and character $(63/n)$ and by Serre–Deligne [11] its Mellin transform is of the form $L(s, \rho)$ for a 2-dimensional dihedral representation ρ. In this case it is induced by the non-trivial character of the Galois group Gal $(Q(\sqrt{21}, \sqrt{-3})Q/(\sqrt{21}))$. [In the statement of (1), $p \equiv 10, 13, 19 \pmod{21}$ is missing].

(4) The Euler product has been determined by Hecke. [2].

(5) This is also found in Hecke and in fact this is the Hecke L-series $L(s, \chi_3, Q(\sqrt{-7}))$, the type we had in § 5.

Acknowledgements

The author is thankful to Professor K G Ramanathan for suggesting this project and for providing him with all the unpublished results of Ramanujan on Dirichlet series with Euler product. He also gratefully acknowledges the help rendered by Mrs. Mythili R Rao for carrying out some of the computations on the TIFR Computer.

References

[1] Birch B J 1975 A look back at Ramanujan's Notebooks ; *Math. Proc. Cambridge Philos. Soc.* **78**, Part I pp. 73–79

[2] Hecke E *Mathematische Werke*, Vandenhoeck and Ruprecht, 1959.
(23) *Zur Theorie der elliptischen Modulfunktionen* (*Math. Ann.* **97** (1926), p. 210–242.
(35, 36) *Uber Modulfunktionen und die Dirichletschen Reihen mit Eulerscher Produktentwicklung.* I, II (*Math. Ann.* **114** (1937), pp. 1–28, pp. 316–351).
(41) *Analytische Arithmetik der positiven quadratischen Formen.* (Kgl. Danske Videnskabernes Selskab. Mathematisk-fysiske Meddelelser, XVII, 12 (1940), 134 pages)].

[3] Li W W 1975 New forms and functional equations ; *Math. Ann.* **212** **4** 285–315

[4] Mordell L J 1916–19 On Mr. Ramanujan's empirical expansions of modular functions ; *Proc. Cambridge Philos. Soc.* **19** (1916–19), pp. 117–124

[5] Newman M 1956 A table of the coefficients of the powers of $\eta(\tau)$ *Indagationes Math.* **18** 204–216

[6] Ramanathan K G 1980 Ramanujan and the congruence properties of partitions : *Proc. Indian Acad. Sci.* (*Math. Sci.*) **89** 133–157

[7] Ramanujan S 1927 *On certain arithmetical functions.* Collected papers of Srinivasa Ramanujan, Cambridge University Press, 136–162

[8] Ramanujan S Unpublished manuscripts.

[9] Rankin R A 1967 Hecke operators on congruence subgroups of the modular group; *Math. Ann.* **168** 40–58

[10] Rankin R A 1977 *Ramanujan's unpublished work on congruences, Modular functions of one variable* V. (eds) J P Serre and D B Zagier Lecture Notes 601 (New York : Springer-Verlag) 3–13

[11] Serre J P *Modular forms of weight one and Galois representations, Durham symposium on algebraic number fields* (L functions and Galois properties) (ed) A Fröhlich, (London : Academic Press) 1977

[12] Serre J P and Stark H M 1977 *Modular forms of weight* $\frac{1}{2}$. Modular functions of one variable VI. (eds) J P Serre and D B Zagier Lecture Notes 627 (New York : Springer Verlag), pp. 27–67

[13] Weber H *Lehrbuch der Algebra*, Vol. III (New York : Chelsea)

Proc. Indian Acad. Sci. (Math. Sci.), Vol. 91, Number 1, January 1982, pp. 17–37.

Oberbeck convection through vertical porous stratum

N RUDRAIAH, M VENKATACHALAPPA and
M S MALASHETTY
UGC–DSA Centre in Fluid Mechanics, Department of Mathematics,
Central College, Bangalore University, Bangalore 560 001, India

MS received 30 November 1981

Abstract. Natural convection through a vertical porous stratum is investigated both analytically and numerically. Analytical solutions are obtained using a perturbation method valid for small values of buoyancy parameter N and the numerical solutions are obtained using Runge-Kutta-Gill method. It is shown that analytical solutions are valid for $N < 1$ and several features of the effect of large values of N are reported. The combined effects of increase in the values of temperature difference between the plates and the permeability parameter on velocity, temperature, mass flow rate and the rate of heat transfer are reported. It is shown that higher temperature difference is required to achieve the mass flow rate in a porous medium equivalent to that of viscous flow.

Keywords. Oberbeck convection; vertical porous stratum; Runge-Kutta-Gill method; buoyancy parameter.

1. Introduction

Density differences in fluid saturated vertical porous medium, interacting with the gravitational field, produces an immense diversity of buoyancy forces and flow configurations in nature and in process of technology; particularly in petroleum and chemical industries. It is also of interest in geohydrology where the flow pattern in heated ground water possesses features of a vertical convection column into which cold ground water is being entrained (see Wooding [7]).

The study of motion of ground water is usually based on the Darcy law [2] in which the macroscopic length scale of a system is so large that the diffusion effects are neglected. However, in zones of mixing between fluids at different temperatures the diffusion effects are significant since the gradients of fluid properties are large. In these zones, the nature of fluid motion is such that boundary layer approximations are valid and the usual potential nature of Darcy equation is not valid. An exact mathematical model to account for the boundary layer aspect in a porous medium is not yet available to our knowledge. To a first approximation, however, we can use the Brinkman model (Rudraiah and Nagaraj [5]). Flows obeying Brinkman model might be found in geothermal areas or might arise from the heat generated by deep explosions in saturated ground.

Therefore, recently Rudraiah and Nagaraj [5] have investigated natural convection in a vertical porous stratum using the Brinkman [1] model and maintaining the boundaries at the same temperature. However, in many practical problems cited above one usually encounters the Oberbeck type of convection (maintaining the temperature gradient in a direction normal to gravity) for which the boundaries have to be maintained at different temperatures; viz., one is maintained at a higher temperature than the other. This Oberbeck problem, in the usual viscous case (that is absence of a porous medium), has been extensively studied (see Gebhart [3]) but much attention has not been given in the case of a porous medium. Recently, Rudraiah and Nagaraj [5] have obtained analytical solutions in a vertical porous layer using a regular perturbation technique which are true only for small values of buoyancy parameter N. It is of interest to know up to what values of N the analytical solutions are valid. These aspects are investigated in this paper using the following plan of work.

The Brinkman equations and the corresponding boundary conditions are given in §2. Analytical solutions for small values of the buoyancy parameter N are obtained in §3. To know up to what values of N the analytical solutions are valid, the basic equations are numerically integrated in §4, using Runge-Kutta-Gill method for wide range of values of N. The effect of permeability and buoyancy parameter N on the skin friction and rate of heat transfer are also calculated. The results are discussed in §5. It is shown that the analytical solutions are valid for values of $N < 1$.

2. Basic equations

The physical model, shown in figure 1, consists of vertical porous stratum bounded by two rigid plates at $y = \pm b$ with x-axis in the axial direction and y-axis perpendicular to the plates. We assume that steady Boussinesq fluid percolating through a homogeneous isotropic porous medium in the x-direction and physical quantities vary with respect to y. Buoyancy forces, due to density difference,

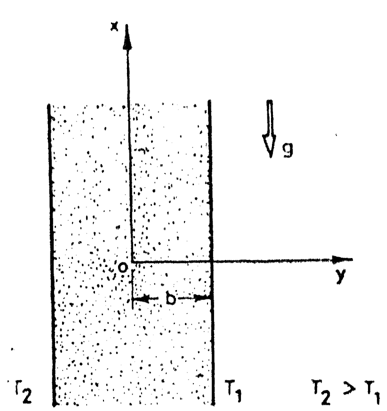


Figure 1. Physical configuration of flow.

cause the fluid to flow upwards through the channel. For this flow, the basic equations of motion, following Rudraiah and Nagaraj [5], are

$$\frac{d^2 u}{dy^2} - \frac{u}{k} + \frac{g\beta (T - T_0)}{\nu} = 0, \tag{1}$$

$$\frac{d^2 T}{dy^2} + \frac{\rho_0 \nu}{K}\left(\frac{du}{dy}\right)^2 + \frac{\rho_0 \nu}{kK} u^2 = 0, \tag{2}$$

$$\rho = \rho_0 [1 - \beta (T - T_0)], \tag{3}$$

where u is the velocity component in the x-direction, T the temperature, T_0 the ambient temperature at $\rho = \rho_0$, ρ the density of the fluid, ν the kinematic viscosity, K the thermal conductivity, k the permeability of a porous medium and β the coefficient of thermal expansion. These equations are solved using the boundary conditions

$$u = 0 \text{ at } y = \pm b, \tag{4}$$

$$T = T_1 \text{ at } y = + b, \tag{5}$$

$$T = T_2 \text{ at } y = - b. \tag{6}$$

The boundary condition on velocity represents the noslip condition and that on temperature points to the fact that the plates are isothermally maintained at different temperatures T_1 and T_2 ($T_2 > T_1$). Equations (1) and (2), using the dimensionless quantities

$$y^* = \frac{y}{b}, \quad \theta = \frac{T - T_0}{T_1 - T_0}, \quad u^* = \frac{\nu}{g\beta b^2 (T_1 - T_0)} u, \tag{7}$$

and for simplicity neglecting the asterisk (*), become

$$\frac{d^2 u}{dy^2} - \sigma^2 u + \theta = 0, \tag{8}$$

$$\frac{d^2 \theta}{dy^2} + N\left(\frac{du}{dy}\right)^2 + N\sigma^2 u^2 = 0, \tag{9}$$

where

$\sigma = b/\sqrt{k}$ is the permeability parameter and

$N = \dfrac{\rho_0 g^2 b^4 \beta^2 (T_1 - T_0)}{\nu K}$ is the buoyancy parameter.

Since the flow is caused by the buoyancy force, the velocity is made dimensionless using this force. The corresponding boundary conditions are

$$u = 0 \quad \text{at} \quad y = \pm 1, \tag{10}$$

$$\theta = 1 \quad \text{at} \quad y = 1, \tag{11}$$

$$\theta = 1 + \bar{\theta} \quad \text{at} \quad y = -1, \tag{12}$$

where

$$\bar{\theta} = \frac{T_2 - T_1}{T_1 - T_0}. \tag{13}$$

At first sight it appears reasonable to ignore the dissipative effects in (9). However, the propagation of waves in a porous medium has small free decay time and hence considerable dissipative effects. Thus the accurate description of flow in a porous medium must include dissipative effects. Therefore (8) and (9) are coupled non-linear equations because of the dissipation term which must be solved simultaneously to yield the desired velocity and temperature profiles. Due to the non-linearity, analytical solutions of these equations are difficult. However if N is small such solutions, following Rudraiah and Nagaraj [5], can be obtained using a regular perturbation technique. This is done in §3, with the motive of understanding the dissipative effects on the flow. To know the validity of these solutions and to find the effects of large N on the flow, (8) and (9) are solved numerically in §4.

3. Analytical solutions

When the buoyancy parameter N is very small we look for solutions of (8) and (9) in the form

$$u = u_0 + Nu_1 + N^2 u_2 + \ldots, \tag{14}$$

$$\theta = \theta_0 + N\theta_1 + N^2\theta_2 + \ldots, \tag{15}$$

where zero subscript quantities are the solutions for the case $N = 0$, which represents physically the solutions in the absence of viscous and Darcy dissipations and $u_1, u_2 \ldots, \theta_1, \theta_2 \ldots$ represents the perturbation quantities relating to u_0 and θ_0. Substituting (14) and (15) into (8)-(12) and equating the coefficients of the like powers of N to zero we get the following set of equations.

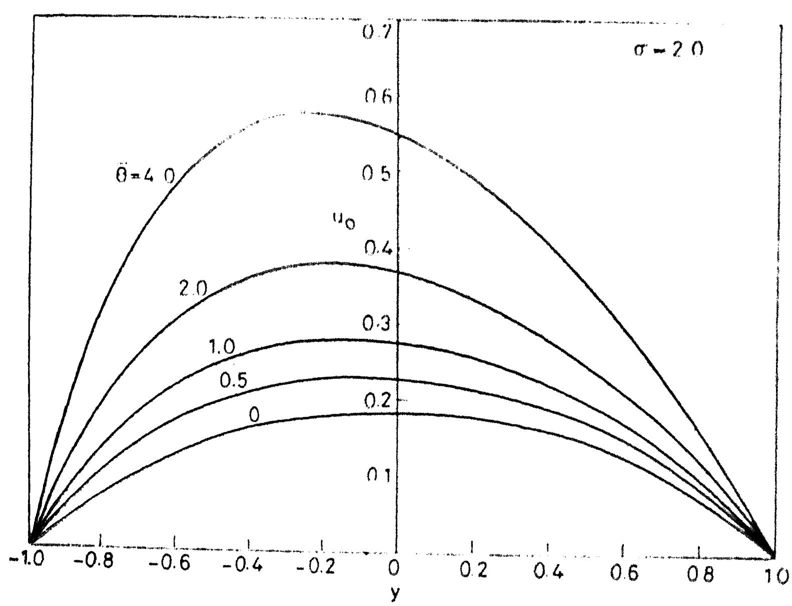


Figure 2. Velocity profiles for various $\bar{\theta}$ (when dissipations are neglected).

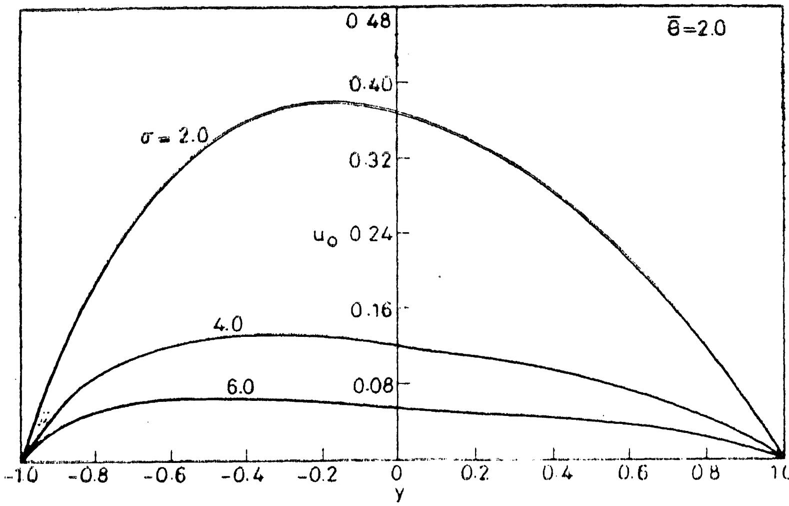


Figure 3. Velocity profiles for various σ (when dissipations are neglected).

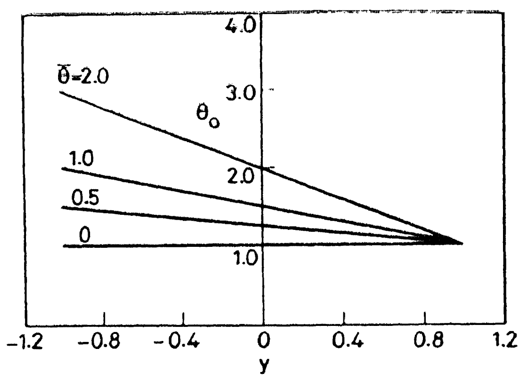


Figure 4. Temperature profiles (when dissipations are neglected).

Zeroth order equations

$$\frac{d^2 u_0}{dy^2} - \sigma^2 u_0 + \theta_0 = 0, \tag{16}$$

$$\frac{d^2 \theta_0}{dy^2} = 0. \tag{17}$$

First order equations

$$\frac{d^2 u_1}{dy^2} - \sigma^2 u_1 + \theta_1 = 0, \tag{18}$$

$$\frac{d^2 \theta_1}{dy^2} + \left(\frac{du_0}{dy}\right)^2 + \sigma^2 u_0^2 = 0, \tag{19}$$

The corresponding boundary conditions are

$$u_0 = 0 \quad \text{at} \quad y = \pm 1, \tag{20}$$

$$\theta_0 = 1 \quad \text{at} \quad y = +1, \tag{21}$$

$$\theta_0 = 1 + \bar{\theta} \quad \text{at } y = -1, \tag{22}$$

and

$$u_i = \theta_i = 0 \text{ at } y = \pm 1, \; i = 1, 2, \ldots \tag{23}$$

The solutions of (16) and (17) using the boundary conditions (20)-(22) are

$$u_0 = \frac{1}{\sigma^2}\left[1 - \frac{\cosh \sigma y}{\cosh \sigma}\right] + \frac{\bar{\theta}}{2\sigma^2}\left[\frac{\sinh \sigma y}{\sinh \sigma} - \frac{\cosh \sigma y}{\cosh \sigma} - y + 1\right], \tag{24}$$

$$\theta_0 = 1 + \frac{\bar{\theta}}{2}\,[1 - y]. \tag{25}$$

In the limit $\sigma \to 0$, (24) tends to

$$u_{f\bullet} = \frac{1}{2}[1 - y^2] + \frac{\bar{\theta}}{12}\,[y^3 - 3y^2 - y + 3], \tag{26}$$

which is the velocity distribution in the channel in the absence of porous material, called viscous flow solutions. When $\bar{\theta} = 0$ that is when the plates are maintained at the same temperature (24) and (25) reduce to those given by Rudraiah and Nagaraj [5]. The second term on the right hand side of (24) and (25) are the contributions from maintaining the plates at different temperatures. Equations (24) and (25) are computed for different values of $\bar{\theta}$ and σ which are shown in figures 2-4 and the results are discussed in §5.

The solutions of (18) and (19), using the boundary conditions (23), are

$$\begin{aligned}
u_1 = {} & f_1(\sigma)\frac{\cosh \sigma y}{\cosh \sigma} - \left[\frac{y}{\sigma^5}\cdot\frac{\sinh \sigma y}{\cosh \sigma} - \frac{\cosh 2\sigma y}{12\sigma^6 \cosh^2 \sigma} + \frac{y^2}{2\sigma^4}\right] \\
& - f_2(\sigma) + \bar{\theta}^2\left[f_3(\sigma)\frac{\cosh \sigma y}{\cosh \sigma} - f_4(\sigma)\frac{\sinh \sigma y}{\sinh \sigma}\right] \\
& + \frac{\bar{\theta}^2}{4\sigma^4}\left\{\frac{\coth 2\sigma}{3\sigma^2}\cdot\frac{\cosh 2\sigma y}{\sinh 2\sigma} + \frac{y(1 - y^2)}{\sigma}\left[\frac{\cosh \sigma y}{\sinh \sigma} - \frac{\sinh \sigma y}{\cosh \sigma}\right]\right. \\
& + \frac{3y^2}{2\sigma^2}\left[\frac{\sinh \sigma y}{\sinh \sigma} - \frac{\cosh \sigma y}{\cosh \sigma}\right] - \frac{\sinh 2\sigma y}{3\sigma^2 \sinh 2\sigma} - \frac{y^4}{12} + \frac{y^3}{3} \\
& \left. + \frac{(\sigma^2 + 2)}{2\sigma^2}\,y^2 + \frac{2y}{\sigma^2} - \frac{1}{\sigma^2} - \frac{2}{\sigma^4}\right\} + \bar{\theta}\left[f_5(\sigma)\frac{\cosh \sigma y}{\cosh \sigma}\right. \\
& \left. - 2f_4(\sigma)\frac{\sinh \sigma y}{\sinh \sigma}\right] + \frac{\bar{\theta}}{\sigma^4}\left\{\frac{\cosh 2\sigma y}{12\sigma^2 \cosh^2 \sigma} - \frac{y \sinh \sigma y}{\sigma \cosh \sigma}\right. \\
& + \frac{y^2}{4\sigma}\cdot\frac{\sinh \sigma y}{\cosh \sigma} + \frac{y}{2\sigma}\cdot\frac{\cosh \sigma y}{\sinh \sigma} - \frac{1}{6\sigma^2}\frac{\sinh 2\sigma y}{\sinh 2\sigma} - \frac{3y \cosh \sigma y}{4\sigma^2 \cosh \sigma} \\
& \left. + \frac{y^3}{6} - \frac{y^2}{2} + \frac{(y - 1)}{\sigma^2}\right\} + f_6(\sigma)\left[1 - \frac{\cosh \sigma y}{\cosh \sigma}\right] \\
& + f_7(\sigma)\left[\frac{\sinh \sigma y}{\sinh \sigma} - y\right],
\end{aligned} \tag{27}$$

$$\theta_1 = \frac{2}{\sigma^4} \cdot \frac{\cosh \sigma y}{\cosh \sigma} - \frac{1}{4\sigma^4} \frac{\cosh 2\sigma y}{\cosh^2 \sigma} - \frac{y^2}{2\sigma^2} + f_8(\sigma)$$

$$- \frac{\bar{\theta}^2}{4\sigma^2}\left[\frac{\cosh 2\sigma}{\sigma^2 \sinh^2 2\sigma} \cdot \cosh 2\sigma y - \frac{\sinh 2\sigma y}{\sigma^2 \sinh 2\sigma} + \frac{2}{\sigma^2}(1-y)\frac{\sinh \sigma y}{\sinh \sigma}\right.$$

$$- \frac{2}{\sigma^2}(1-y)\frac{\cosh \sigma y}{\cosh \sigma} + \frac{2}{\sigma^3}\left(\frac{\cosh \sigma y}{\sinh \sigma} - \frac{\sinh \sigma y}{\cosh \sigma}\right) + \frac{y^4}{12}$$

$$\left. - \frac{y^3}{3} + \frac{y^2}{2}\right] + \frac{\bar{\theta}}{\sigma^2}\left[\frac{1}{\sigma^2}\left(\frac{\sinh \sigma y}{\sinh \sigma} - 2\frac{\cosh \sigma y}{\cosh \sigma}\right)\right.$$

$$- \frac{\sinh 2\sigma y}{2\sigma^2 \sinh 2\sigma} + \frac{\cosh 2\sigma y}{4\sigma^2 \cosh^2 \sigma} + \frac{y \cosh \sigma y}{\sigma^2 \cosh \sigma} - \frac{1}{\sigma^3}\frac{\sinh \sigma y}{\cosh \sigma}$$

$$\left. - \frac{y^3}{6} + \frac{y^2}{2}\right] + f_6(\sigma) \cdot y + f_7(\sigma), \tag{28}$$

where

$$f_1(\sigma) = \frac{3}{\sigma^6} + \frac{\tanh \sigma}{\sigma^5} - \frac{\cosh 2\sigma}{3\sigma^6 \cosh^2 \sigma},$$

$$f_2(\sigma) = \frac{3}{\sigma^6} - \frac{\cosh 2\sigma}{4\sigma^6 \cosh^2 \sigma} - \frac{1}{2\sigma^4},$$

$$f_3(\sigma) = \frac{1}{4\sigma^4}\left[-\frac{\coth^2 2\sigma}{3\sigma^2} + \frac{\coth \sigma}{2\sigma} + \frac{\tanh \sigma}{\sigma} + \frac{2}{\sigma^4} + \frac{1}{2\sigma^2} + \frac{7}{12}\right],$$

$$f_4(\sigma) = \frac{1}{4\sigma^4}\left[\frac{\coth \sigma}{\sigma} + \frac{\tanh \sigma}{2\sigma} + \frac{1}{6\sigma^2} + \frac{1}{3}\right],$$

$$f_5(\sigma) = \frac{1}{\sigma^4}\left[-\frac{\cosh 2\sigma}{12\sigma^2 \cosh^2 \sigma} + \frac{\tanh \sigma}{\sigma} + \frac{1}{\sigma^2} + \frac{1}{2}\right],$$

$$f_6(\sigma) = \frac{\bar{\theta}^2}{4\sigma^2}\left[\frac{\cosh^2 2\sigma}{\sigma^2 \sinh^2 2\sigma} + 2\frac{\coth \sigma}{\sigma^3} - \frac{4}{\sigma^2} + \frac{7}{12}\right]$$

$$+ \frac{\bar{\theta}}{\sigma^2}\left[\frac{\cosh 2\sigma}{4\sigma^2 \cosh^2 \sigma} - \frac{2}{\sigma^2} + \frac{1}{2}\right],$$

$$f_7(\sigma) = \frac{\bar{\theta}(\bar{\theta}+2)}{2\sigma^2}\left[\frac{\tanh \sigma}{\sigma^3} - \frac{3}{2\sigma^2} + \frac{1}{6}\right],$$

and

$$f_8(\sigma) = \frac{1}{4\sigma^4} \cdot \frac{\cosh 2\sigma}{\cosh^2 \sigma} - \frac{2}{\sigma^4} + \frac{1}{2\sigma^2}.$$

Equations (27) and (28) are evaluated for different values of σ and $\bar{\theta}$ and the results are shown in figures 5 and 6.

It is of practical interest to find the mass flow rate and friction factor. The mass flow rate depicts quantitatively the effect of permeability on the flow and the friction factor gives information as to under what Reynolds number the flow is laminar.

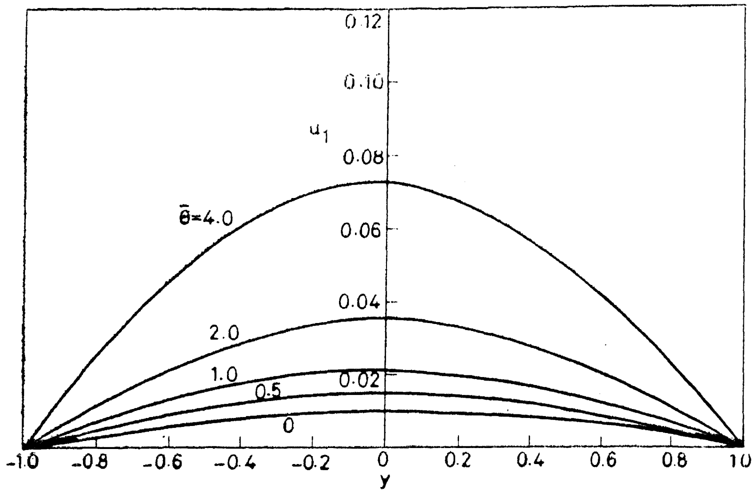


Figure 5. Velocity correction profiles for different $\bar{\theta}$ and $\sigma = 2\cdot 0$.

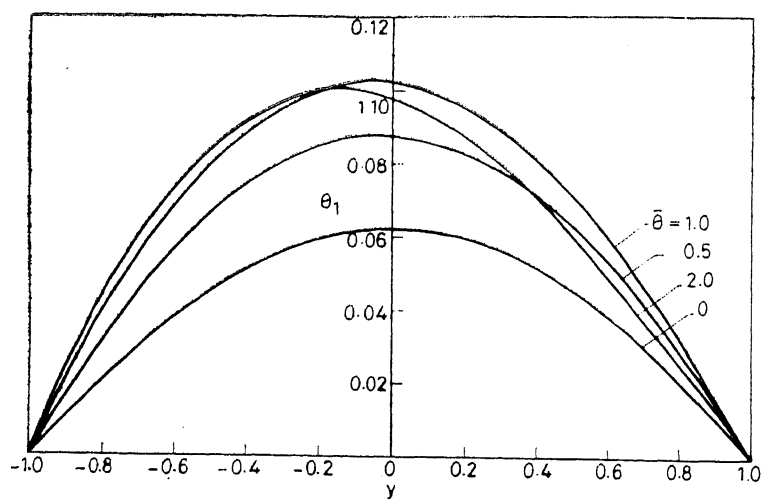


Figure 6. Temperature correction profiles for different $\bar{\theta}$ and $\sigma = 2\cdot 0$.

If M_{p_0} denotes the mass flow rate per unit channel width in the absence of dissipation (i.e., $N = 0$) then

$$M_{p_0} = \int_{-b}^{+b} \rho_0 u_0 dy,$$

$$= -\frac{\rho_0 G b^3}{\sigma^2}(2 + \bar{\theta})\left[1 - \frac{\tanh \sigma}{\sigma}\right], \tag{29}$$

where

$$G = -\frac{g\beta(T_1 - T_0)}{\nu}.$$

On the other hand if M_{f0} denotes the mass flow rate per unit channel width for viscous flow then

$$M_{f_0} = -\frac{\rho_0 G b^2}{3}(2 + \bar{\theta}). \tag{30}$$

For the conditions of equal pressure gradients and channel width the ratio of equation (29) to (30) gives

$$\frac{M_{p_0}}{M_{f_0}} = \frac{3}{\sigma^2}\left[1 - \frac{\tanh \sigma}{\sigma}\right] \tag{31}$$

which is independent of $\bar{\theta}$ because of the neglection of dissipative effects. This ratio of mass flow rates is computed for different values of σ and is shown in figure 7. From this it is clear that mass flow rate decreases with increase in σ.

Now the friction factor C_f is defined by

$$C_f = -\frac{\nu G D}{\frac{1}{2}\bar{u}_0^2}, \tag{32}$$

where D is the equivalent diameter ($D = 4b$ for the channel considered in this paper).

Since

$$\bar{u}_0 = \frac{1}{2b}\int_{-b}^{+b} u_0\, dy,$$

$$= -\frac{Gb^2}{2\sigma^2}(2 + \bar{\theta})\left[1 - \frac{\tanh \sigma}{\sigma}\right], \tag{33}$$

we get

$$C_f = \frac{64\sigma^2}{\mathrm{Re}\,(2 + \bar{\theta})\left[1 - \frac{\tanh \sigma}{\sigma}\right]}, \tag{34}$$

where

$\mathrm{Re} = \frac{u_0 D}{\nu}$ is the Reynolds number.

Thus

$$C_f\,\mathrm{Re} = \frac{64\sigma^2}{(2 + \bar{\theta})\left[1 - \frac{\tanh \sigma}{\sigma}\right]}. \tag{35}$$

This shows that the product $C_f\mathrm{Re}$ is constant (independent of a Reynolds number) for a channel of fixed width and a given porous medium. This product, in the case of usual viscous flow through a channel, is

$$C_f^*\,\mathrm{Re}^* = \frac{192}{(2 + \bar{\theta})}. \tag{36}$$

The ratio of (35) to (36) is

$$\frac{C_f\,\mathrm{Re}}{C_f^*\,\mathrm{Re}^*} = \frac{\sigma^2}{3\,[1 - \tanh \sigma/\sigma]}, \tag{37}$$

which depends only on the permeability of a porous medium and the width of the channel. This ratio is numerically evaluated for different values of σ and is shown in figure 8. From this it is clear that the ratio increases with increase in σ. Thus for large value of σ the flow is laminar.

Once the velocity distribution is known, the skin friction can be calculated from

$$\tau' = \mu \left(\frac{\partial u}{\partial y}\right)_{y=\pm b}, \tag{38}$$

which in the non-dimensional form can be written, by using (7), as

$$\begin{aligned} \tau &= \frac{\tau'}{\rho g b (T_1 - T_0)}, \\ &= \left(\frac{du}{dy}\right)_{y=\pm 1}, \\ &= \left(\frac{du_0}{dy} + N\frac{du_1}{dy}\right)_{y=\pm 1}. \end{aligned} \tag{39}$$

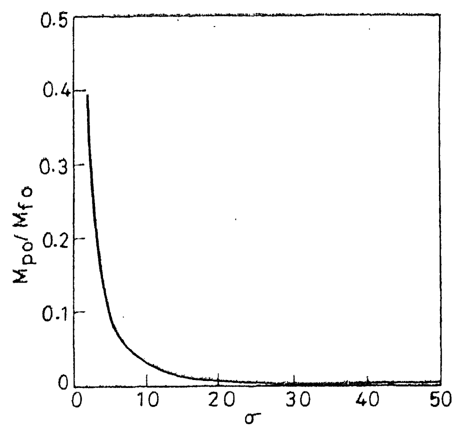


Figure 7. Ratio of mass flow rate with and without porous medium *versus* σ.

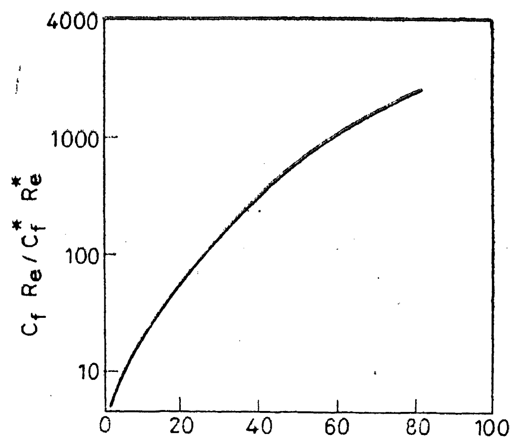


Figure 8. Ratio of the product of friction factor and Reynolds number with and without porous medium *versus* σ.

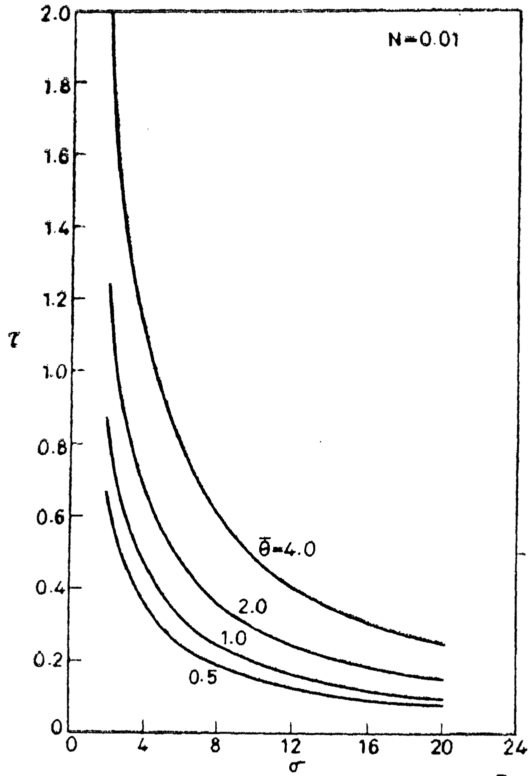


Figure 9. Skin friction *versus* σ for different values of $\bar{\theta}$ at hotter plate.

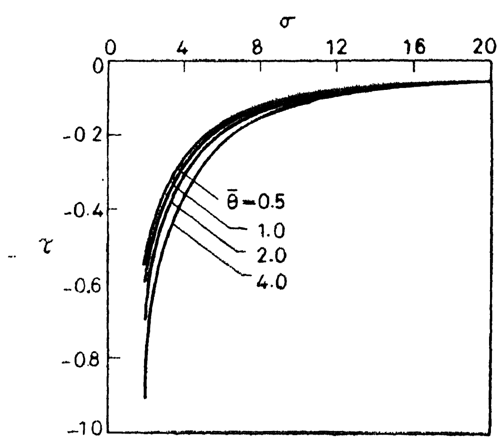


Figure 10. Skin friction *versus* σ for different values of $\bar{\theta}$ at cooler plate.

τ is plotted against σ in figures 9 and 10, for $N = 0{\cdot}01$ and for different values of $\bar{\theta}$.

It is of practical interest and importance to calculate the rate of heat transfer between the fluid and the plates. This is given by

$$q' = -K\left(\frac{\partial T}{\partial y}\right)_{y=\pm b}, \tag{40}$$

which in the non-dimensional form can be written using (7) as

$$q = \frac{-q'b}{K(T_1 - T_0)},$$

$$= \left(\frac{d\theta}{dy}\right)_{y=\pm 1}, \tag{41}$$

$$= \left(\frac{d\theta_0}{dy} + N\frac{d\theta_1}{dy}\right)_{y=\pm 1}. \tag{42}$$

q is plotted against σ in figures 11 and 12 for $N = 0{\cdot}01$ and for different values of $\bar{\theta}$.

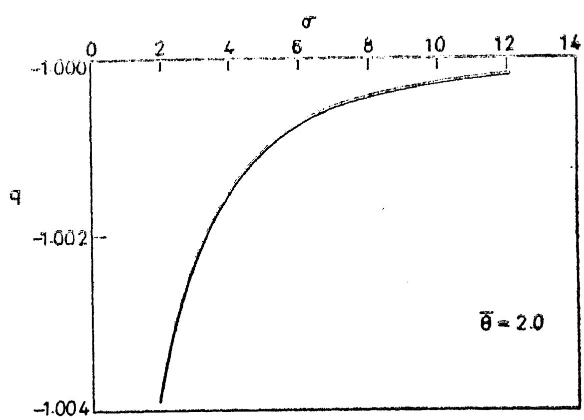


Figure 11. Rate of heat transfer for various σ at hotter plate.

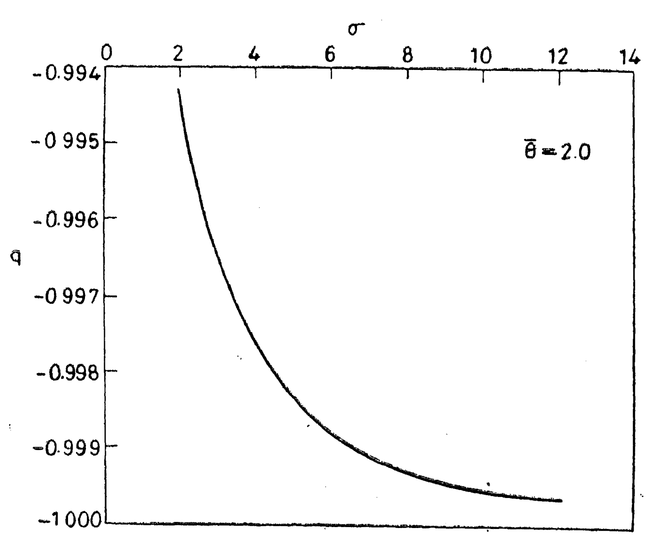


Figure 12. Rate of heat transfer for various σ at cooler plate.

4. Numerical solutions

The analytical solutions obtained in §3 are valid only for small values of the buoyancy parameter N. To know the validity of the analytical solutions and to find the effect of large N on the flow we have solved (8) and (9) with the boundary conditions (10)–(12) numerically, using Runge–Kutta–Gill method.

The numerical method involves solving a non-linear two point boundary value problem using an iterative scheme. We have used a slightly modified method and the FORTRAN program developed by Sen and Venkataramudu [6]. The velocity and temperature distributions are obtained for a wide range of values of N and are shown in figures 13 and 14. We find that the analytical solutions are in very good agreement with the numerical solutions when N is very small. For the sake of comparison we have given the values of velocity and temperature at $y=0$ for $\bar{\theta}=1\cdot0$ and $\sigma=2\cdot0$ in tables 1 and 2.

The mass flow rate and friction factor found in §3 are true only in the absence of dissipations. Here we find the mass flow rate and friction factor in the presence of dissipative effects.

Let M_p denote the mass flow rate per unit channel width in the presence of dissipations, then

$$M_p = \int_{-b}^{+b} \rho_0 u dy,$$

$$= -\rho_0 G b^3 \int_{-1}^{+1} u dy. \tag{43}$$

Let M_f denote the mass flow rate per unit channel width in the presence of dissipation for viscous flow, then

$$M_f = -\rho_0 G b^3 \int_{-1}^{+1} u_f dy. \tag{44}$$

The ratio of (43) to (44) is

$$\frac{M_p}{M_f} = \frac{\int_{-1}^{+1} u dy}{\int_{-1}^{+1} u_f dy}. \tag{45}$$

The friction factor C_f is defined as

$$C_f = -\frac{\nu G D}{\frac{1}{2}\bar{u}^2}, \tag{46}$$

where

$$\bar{u} = \frac{1}{2b}\int_{-b}^{+b} u dy,$$

$$= -\frac{Gb^2}{2}\int_{-1}^{+1} u dy. \tag{47}$$

Equation (46) using (47) becomes

$$C_f = \frac{64}{\mathrm{Re}\int_{-1}^{+1} u dy}, \tag{48}$$

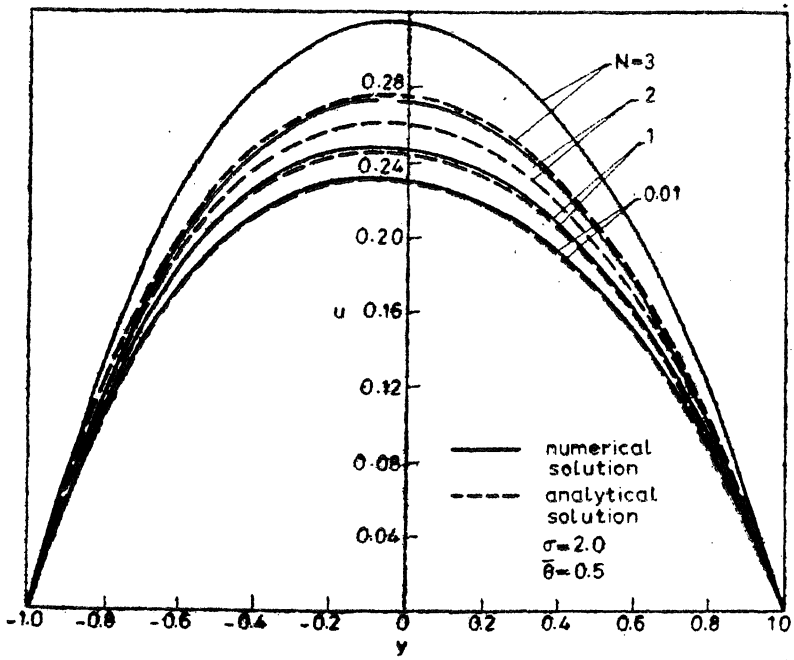


Figure 13. Velocity profiles.

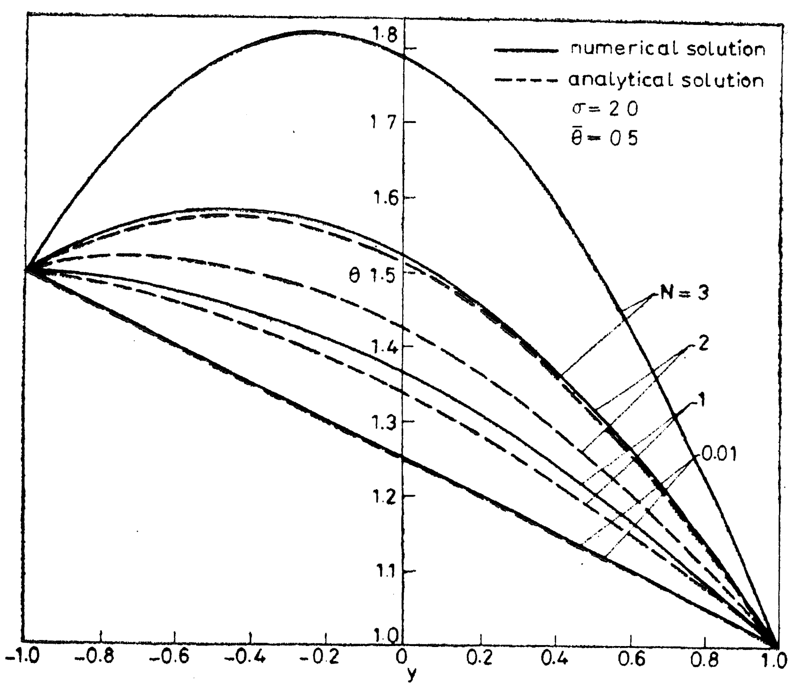


Figure 14. Temperature profiles.

Table 1. Velocity at $y = 0$, $\bar{\theta} = 1\cdot0$, $\sigma = 2\cdot0$.

N	Numerical	Analytical
0·01	0·2755473	0·2755308
0·05	0·2764471	0·2763728
0·1	0·2778859	0·2774250
0·2	0·2790284	0·279530
0·3	0·2823414	0·2816350
0·4	0·2846408	0·283740
0·5	0·2874295	0·285845
0·6	0·2901138	0·287950
0·7	0·2928998	0·290055
0·8	0·2957960	0·292160
0·9	0·2988109	0·294265

Table 2. Temperature at $y = 0$, $\bar{\theta} = 1\cdot0$, $\sigma = 2\cdot0$.

N	Numerical	Analytical
0·01	1·501418	1·501028
0·05	1·507135	1·5051260
0·1	1·514384	1·510252
0·2	1·529239	1·520504
0·3	1·544593	1·530756
0·4	1·560480	1·541008
0·5	1·576935	1·551260
0·6	1·593999	1·561512
0·7	1·611715	1·571764
0·8	1·630132	1·582016
0·9	1·649305	1·592268

where

$$\mathrm{Re} = \frac{\bar{u}D}{\nu} \text{ is the Reynolds number.}$$

Thus

$$C_f\,\mathrm{Re} = \frac{64}{\int_{-1}^{+1} u\,dy}. \tag{49}$$

This product in the case of viscous flow is

$$C_f{}^*\,\mathrm{Re}^* = \frac{64}{\int_{-1}^{+1} u_f\,dy} \tag{50}$$

The ratio of (49) to (50) is

$$\frac{C_f \, Re}{C_f^* \, R_e^*} = \frac{\int_{-1}^{+1} u_f \, dy}{\int_{-1}^{+1} u dy}. \tag{51}$$

Equations (45) and (51) are evaluated numerically for different values of N and are shown in figures 15 to 17.

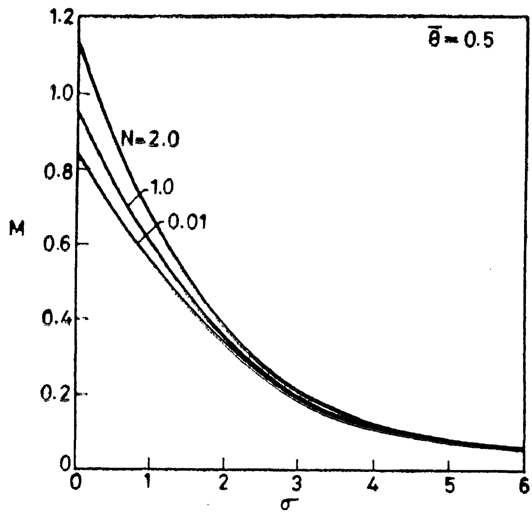


Figure 15. Mass flow rate *vs.* σ for different N.

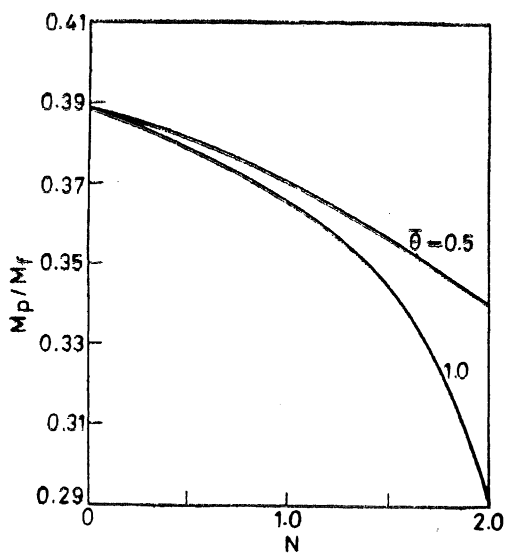


Figure 16. Ratio of mass flow rate *vs.* N for $\sigma = 2 \cdot 0$.

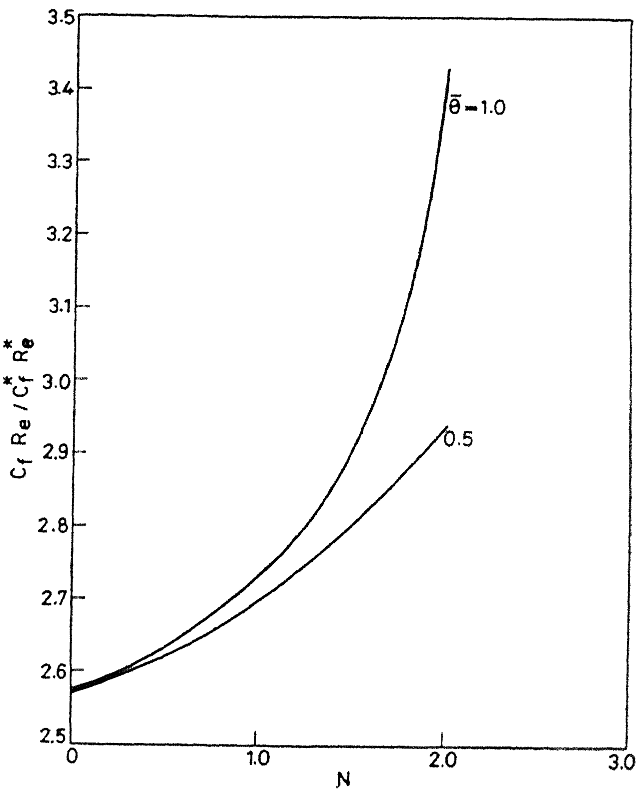


Figure 17. Ratio of the product of friction factor and Reynolds number *vs. N*.

The skin friction and the rate of heat transfer defined by (38) and (40) are numerically evaluated for different values of N and σ and are shown in figures 18 to 21.

5. Discussion

The direct analytical solutions for the problem of natural convection through a vertical porous stratum can be obtained in the absence of dissipative effects. In the presence of dissipative effects, however, analytical solutions are obtained by a regular perturbation method valid for small values of N and the results are shown in figures 2 to 12. Figures 2 to 4 are concerned with velocity and temperature distributions in the absence of dissipative effects. The effect of increase in the temperature difference between the plates, viz., $\bar{\theta}$, is to increase the velocity and temperature distributions (see figures 2 and 4) due to increase in convection. The perturbation in velocity and temperature due to dissipation is small which is evident from figures 5 and 6. The effect of increase in σ is to decrease the velocity and temperature distributions because of the dampening effect of the Darcy resistance.

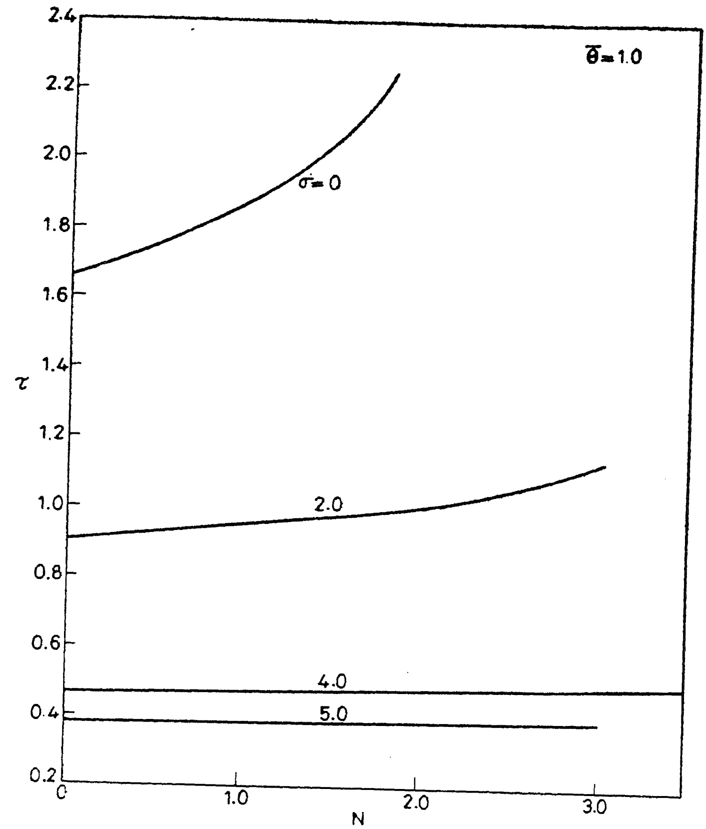


Figure 18. Skin friction *vs.* N at hotter plate.

The ratio of mass flow rate and friction factor with and without porous media, in the absence of dissipative effects, are shown in figures 7 and 8. We see that as σ increases the mass flow rate decreases and friction factor increases. This decrease in mass flow rate with increase in σ is very useful in studying the pore size distribution in a porous medium. The increase in friction factor with increase in σ ensures the laminar flow. The skin friction and rate of heat transfer are computed in the absence of disspative effects for different values of σ and $\bar{\theta}$ and the results are shown in figures 9 to 12. The skin friction increases with increase in $\bar{\theta}$ and decreases with increasing σ. The rate of heat transfer decreases numerically near the hotter plate and increases near the cooler plate with an increase in σ which is evident from the physical grounds explained earlier.

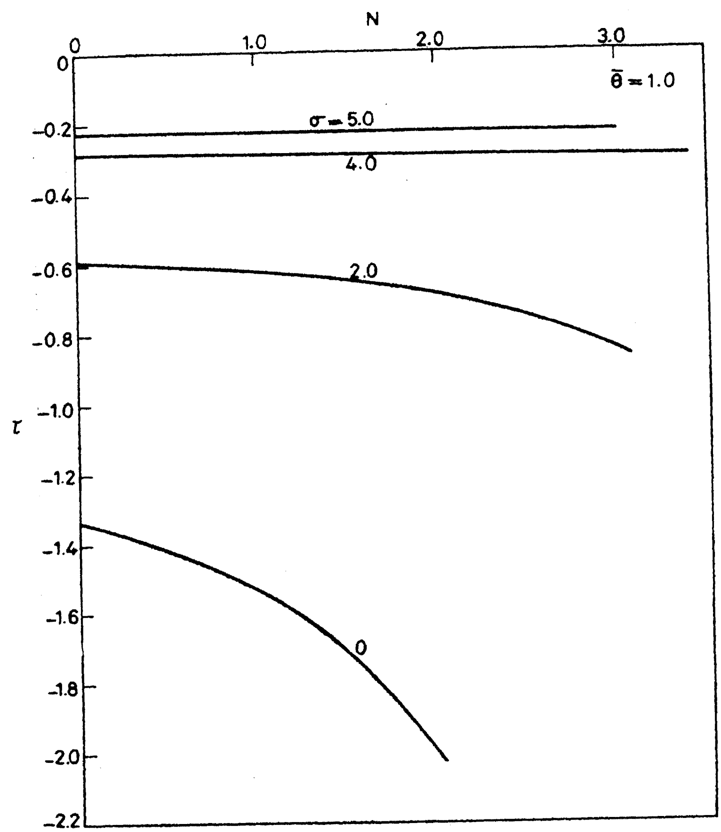


Figure 19. Skin friction *vs.* N at cooler plate.

The above results obtained by the pert urbation method are true only for small values of $N(<1)$. To know up to what values of N the perturbation solutions are valid and to find the effect of large N on the flow, we have solved (8) and (9) numerically using Runge–Kutta–Gill method. The velocity and temperature distributions are obtained for a wide range of values of N and the results are compared with the perturbation solutions in figures 13 and 14. We find that the analytical solutions are in good agreement with the numerical results up to $N = 1$ and they deviate considerably for $N > 1$. We also observe that the velocity and temperature distribution, increase with increase in N with higher values near the hotter plate than those at the cooler one.

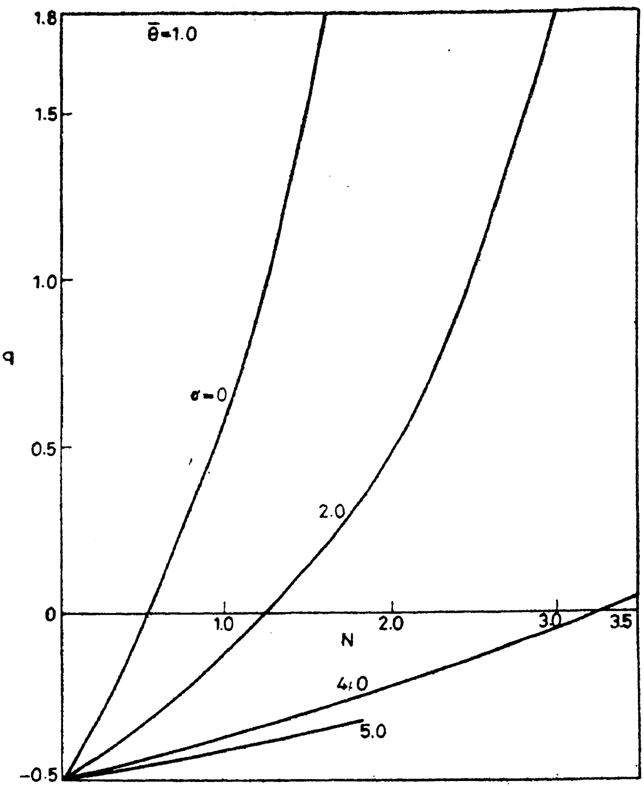


Figure 20. Rate of heat transfer *vs.* N at hotter plate.

The mass flow rate and friction factor in the presence of dissipative effects are computed using the numerical method and the results are shown in figures 15 to 17. From figure 15 it is clear that the increase in N increases the mass flow rate. Although M_p and M_f increase individually with N, their ratio M_p/M_f decreases with N as shown in figure 16. This means that the rate of increase in M_f is much higher than that of M_p. The ratio of friction factor increases with increase in N as shown in figure 17. The skin friction and rate of heat transfer are also computed using the numerical method and the results are shown in figures 18 to 22. We observe that for small values of σ the skin friction increases with N and is independent of N for large values of σ. This is because for small values of σ there exists Taylor-porous boundary layer near the surface in which the velocity gradient is fairly large. For large values of σ the boundary layer type of flow governed by Brinkman model transforms to potential nature of Darcy flow (see Rudraiah and Masuoka [4]) where the velocity gradients are negligibly small. The rate of heat transfer, as shown in figures 20 and 21, increases numerically with N. This is because as N increases the temperature difference also increases resulting in large convection which transfers more heat.

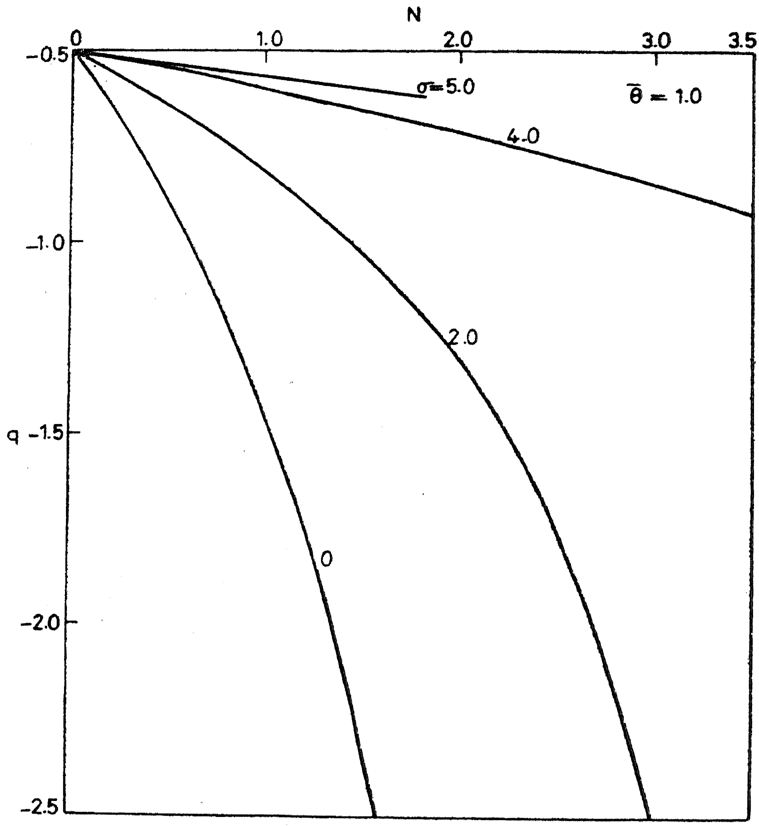


Figure 21. Rate of heat transfer *vs.* N at cooler plate.

Acknowledgements

This work was supported by the UGC India under the DSA scheme. The work of one of the authors (MSM) is supported by the UGC India under its faculty improvement programme. MSM is also grateful to HKE Society for deputing him for higher studies.

References

[1] Brinkman H C 1947 *Appl. Sci. Res.* **A1** 27

[2] Darcy H 1956 Les Fontaines publiques de la Ville de Dijon (Paris : Victor Dalmint)

[3] Gebhart B 1979 *ASME J. Fluid Eng.* **101** 5

[4] Rudraiah N and Masuoka T 1982 *Int. J. Eng. Sci.* **20** 27

[5] Rudraiah N and Nagaraj S T 1977 *Int. J. Eng. Sci.* **15** 589

[6] Sen S K and Venkataramudu V 1976 Report No. CC/SKS-VV/R-04-76 Indian Institute of Science, Bangalore

[7] Wooding R A 1960 *J. Fluid Mech.* **7** 501

Proc. Indian Acad. Sci. (Math. Sci.), Vol. 91, Number 1, January 1982, pp. 39–51.

Long waves in inviscid compressible atmosphere II

P L SACHDEV and V S SESHADRI

Department of Applied Mathematics, Indian Institute of Science, Bangalore 560 012, India

MS received 4 June 1980 ; revised 16 December 1981

Abstract. Solitary waves have been found in an adiabatic compressible atmosphere which, in ambient state, has winds and temperature gradient, generalizing our earlier results for the isothermal atmosphere. Explicit results are obtained for the special case of linear temperature and linear wind distributions in the undisturbed conditions. An important result of the study is that the number of possible critical speeds of the flow depends crucially on whether the maximum Richardson number (which is variable in the present example) is greater or less than 1/4.

1. Introduction

In a previous paper, (Sachdev and Seshadri [3], I hereafter) we derived the equations governing solitary and other long waves in an isothermal atmosphere with wind shear. We discussed the existence and cardinality of solitary waves *via* an eigenvalue problem for a second order ordinary differential equation. In the present paper we generalize these results to more general atmospheres.

The plan of this paper is as follows. In § 2 we formulate our problem for some general atmospheres and discuss the zeroth and first order approximations. In § 3 we specialise the results of § 2 to the case of a linearly increasing temperature profile and derive a model equation. In § 4, the results of § 3 are applied to linearly increasing wind profile. Finally, we give the conclusions of our study in § 5.

2. Formulation and first order analysis

The more general atmosphere is again assumed to be adiabatic and inviscid, extending from the plane $y = 0$ to infinity in the vertical direction y. In its equilibrium state, it is assumed to have pressure and density distributions that are strictly decreasing functions of y, tending to zero as $y \to \infty$ in such a manner that the absolute temperature is a non-decreasing function of y. Further, an ambient shear flow is also assumed to exist, given by $u_0(y)$, the component of velocity in the horizontal direction.

The equations governing the propagation of atmospheric gravity waves are

$$\frac{\partial \rho}{\partial t} + \frac{\partial (\rho u)}{\partial x} + \frac{\partial (\rho v)}{\partial y} = 0, \tag{1a}$$

$$\rho\left(\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}\right)=-\frac{\partial p}{\partial x}, \tag{1b}$$

$$\rho\left(\frac{\partial v}{\partial t}+u\frac{\partial v}{\partial x}+v\frac{\partial v}{\partial y}\right)=-\frac{\partial p}{\partial y}-\rho g, \tag{1c}$$

$$\rho\left(\frac{\partial p}{\partial t}+u\frac{\partial p}{\partial x}+v\frac{\partial p}{\partial y}\right)-\gamma p\left(\frac{\partial \rho}{\partial t}+u\frac{\partial \rho}{\partial x}+v\frac{\partial \rho}{\partial y}\right)=0, \tag{1d}$$

where p is the pressure, ρ is the density, (u, v) is the velocity vector, and g is the acceleration due to gravity.

A wave of permanent form moving with a velocity $(c, 0)$ is assumed to have been generated by some disturbance. This, for example, could be the result of an unsteady motion created much earlier. It is known from the theory of water waves that the propagation speed c of long water waves of permanent form is close to some critical value c^*. The purpose of the present analysis is to find the critical speeds of these waves for the model under study and also the equations governing their propagation.

As in I, equation (1) may be rendered non-dimensional by

$$X = x/H,\ Y = y/H,\ T = tc/H,\ R = \rho/\rho_g.$$
$$P = p/(\rho_g c^2),\ U = u/c,\ V = v/c\ ;$$

the small parameter for the present problem is provided by

$$\epsilon = L - \lambda, \tag{2}$$

where

$$L = (gH/c^2),\ \text{and}\ \lambda = (gH/c^{*2}),$$

and the stretchings are defined by

$$\xi = \epsilon^{1/2}(X - T), \tag{3a}$$

$$\bar{V} = \epsilon^{1/2}V,\ \bar{U} = U - 1. \tag{3b}$$

The notations here are the same as in I.

In terms of the new variables the undisturbed conditions, denoted by the suffix e, satisfy

$$\bar{U}_e(Y) = [u_0(YH)/c - 1],\ \bar{V}_e(Y) = 0, \tag{4}$$

$$\frac{dp_e}{dY} = -\lambda R_e(Y). \tag{5}$$

The governing equations for $\bar{U}, \bar{V}, R$ and P, the boundary conditions at the ground and the perturbation scheme for these functions are the same as 7 (a)-7 (d), (8), (9) and (10) in I.

To the zeroth order, we obtain

$$\bar{U}_0(Y) = \bar{U}_e(Y),\ R_0(Y) = R_e(Y), \tag{6}$$

while $P_0(Y)$ is obtained by solving

$$\frac{dP_0}{dY} = -LR_0(Y), \tag{7}$$

with the initial condition $P_0(0) = L$. By our assumption, $P_0(Y)$ and $R_0(Y)$ are strictly monotonic functions of Y, which decrease to zero as $Y \to \infty$, and the ratio P_0/R_0, which is proportional to the absolute temperature, is a nondecreasing function of Y. The equations for the first order approximations and the boundary conditions are again given by (12a)–(12d) and (13) in I.

Besides,

$$\bar{U}_1(\pm\infty, Y) = 0,\ \bar{V}_1(\pm\infty, Y) = 0,$$
$$P_1(\pm\infty, Y) = P_{e_1}(Y),\ R_1(\pm\infty, Y) = 0, \tag{8}$$

where $P_{e_1}(Y)$ is obtained by solving

$$\frac{dP_{e_1}}{dY} = R_{e0}(Y), \tag{9}$$

with the initial condition $P_{e_1}(0) = -1$.

Eliminating $\bar{U}_1$, $\bar{V}_1$, and R_1 from their perturbation equations we get a single differential equation for P_1:

$$\frac{\partial}{\partial Y}\left[\left(R_0\frac{\partial P_1}{\partial \xi} + \frac{\gamma P_0}{L}\frac{\partial^2 P_1}{\partial \xi\,\partial Y}\right)\Big/\left(\frac{dP_0}{dY} - \frac{\gamma P_0}{R_0}\frac{dR_0}{dY}\right)\right]$$
$$+ \frac{1}{L}\frac{\partial^2 P_1}{\partial \xi\,\partial Y} + \frac{1}{\bar{U}_0^2}\frac{\partial P_1}{\partial \xi} = 0. \tag{10}$$

We exclude critical levels occurring in the flow region, the function $\bar{U}_0(Y) = [u_0(YH)/c - 1]$ is therefore non-zero for $Y \in [0, \infty)$. The factor $[dP_0/dY - (\gamma P_0/R_0)\,dR_0/dY]$ is also not equal to zero for $Y \in [0, \infty)$ by our assumptions on $P_0(Y)$ and $R_0(Y)$.

The boundary condition $\bar{V}_1(\xi, 0) = 0$ can be shewn to be equivalent to

$$\frac{\gamma P_0}{LR_0}\frac{\partial^2 P_1}{\partial \xi\,\partial Y} + \frac{\partial P_1}{\partial \xi} = 0 \text{ at } Y = 0. \tag{11}$$

Further, we impose the condition that the first order terms do not grow larger in amplitude, compared to their zeroth order counterparts. This requires

$$\left|\frac{P_1}{P_0}\right| = 0\,(1) \text{ for all } Y. \tag{12}$$

Equation (10) can be transformed to

$$\frac{\partial^2 \bar{H}}{\partial \eta^2} + \left[LR_0\Big/\left(\gamma \bar{U}_0\frac{d\eta}{dY}\right)\right]\bar{H} = 0, \tag{13}$$

where

$$\bar{H} = \frac{1}{R_0}\frac{\partial P_1}{\partial \xi}, \tag{14}$$

and

$$\eta = \int_0^Y \left(\frac{1}{R_0 P_0}\frac{dP_0}{dY'} - \frac{\gamma}{R_0^2}\frac{dR_0}{dY'}\right) dY'. \tag{15}$$

If T denotes the absolute temperature normalized by $p_g/(\rho_g R_g)$ where R_g is the universal gas constant, we have

$$P/R = \lambda T. \tag{16}$$

If T_0 denotes the normalised temperature to the zeroth order, we have

$$\eta = \int_0^Y \frac{1}{R_0}\left[\frac{1}{T_0}\frac{dT_0}{dY'} - \frac{(\gamma-1)}{R_0}\frac{dR_0}{dY'}\right] dY'. \tag{17}$$

By our assumptions on $P_0(Y)$, $R_0(Y)$ and $T_0(Y)$, (17) gives

$$\eta(Y) > -(\gamma-1)\int_0^Y \frac{1}{R_0^2}\frac{dR_0}{dY'}\,dY' = (\gamma-1)\left[\frac{1}{R_0(Y)} - 1\right]. \tag{18}$$

This implies that $\eta \to \infty$ as $Y \to \infty$. Moreover, the integrand in (17) is strictly positive. We, therefore, conclude that the transformation (15) is a one-one mapping of $Y \in [0, \infty)$ onto $\eta \in [0, \infty)$.

Applying a theorem due to Hille and Wintner [2], we find that (13) has a solution $H_1(\eta)$, unique up to a multiplicative constant, such that

$$H_1(\eta) = 0(1),\ H_1'(\eta) = 0(1/\eta) \text{ as } \eta \to \infty, \tag{19}$$

if and only if

$$\int_0^\infty \left[R_0\eta\Big/\left(\bar{U}_0^2 \frac{d\eta}{dY}\right)\right] d\eta < \infty,$$

that is,

$$\int_0^\infty [R_0\eta/\bar{U}_0^2(Y)]\, dY < \infty, \tag{20}$$

where η is given by (17). There is also another solution $H_2(\eta)$, non-unique, such that

$$H_2(\eta) = 0(\eta),\ H_2'(\eta) = 0(1) \text{ as } \eta \to \infty. \tag{21}$$

Assuming that (20) holds, we obtain

$$\frac{1}{R_0}\frac{\partial P_1}{\partial \xi} = \frac{\partial A_1}{\partial \xi} H_1(\eta) + \frac{\partial A_2}{\partial \xi} H_2(\eta), \tag{22}$$

where $A_1'(\xi)$ and $A_2'(\xi)$ are to be found from the boundary conditions. Integrating (22) with respect to ξ and using the conditions at $|\xi| = \infty$, we get.

$$\frac{P_1}{R_0} = [A_1(\xi) - A_1(\infty)]\, H_1(\eta) + [B_1(\xi) - B_1(\infty)]\, H_2(\eta) + \frac{P_{1_1}}{R_0}. \tag{23}$$

To satisfy condition (12), we put $[B_1(\xi) - B_1(\infty)] = 0$.

Assuming $A_1(\infty) = 0$, we get

$$P_1 = R_0(Y)\, A_1(\xi)\, H_1(\eta) + P_{1*}(Y). \tag{24}$$

From the asymptotic behaviour of $H_1(\eta)$, given in (19), which arises as a result of imposing the condition (20) on the function $\bar{U}_0(Y)$, the right hand side of (24) is $0(R_0)$ as $Y \to \infty$. To satisfy the condition (12), however, it is sufficient that it is $0(P_0)$ as $Y \to \infty$. Since (P_0/R_0), which is proportional to the temperature, is assumed to be a non-decreasing function of Y we conclude that the condition (20) is only a sufficient condition for (12) to be true. For the isothermal case given in I, however, the condition (20) is also necessary for (12) to be satisfied. In this case it is easy to see that

$$\eta = (\gamma - 1)\,[\exp(Y) - 1]$$

and therefore (20) becomes

$$\int_0^\infty \frac{dY}{\bar{U}_0^2(Y)} < \infty,$$

which is the same as the condition (21) of I.

Once the expression for P_1 has been obtained, those for R_1, $\bar{V}_1$, $\bar{U}_1$ can be obtained from first order perturbation equations and the equilibrium conditions (4) and (5).

The vanishing of $\bar{V}_1(\xi, 0)$ gives

$$\frac{\partial P_1}{\partial \xi} + \frac{\gamma P_0}{L R_0}\frac{\partial^2 P_1}{\partial \xi\, \partial Y} = 0 \quad \text{at} \quad Y = 0.$$

This, with the help of (24), becomes

$$H_1(0) - \gamma H_1'(0) = 0. \tag{25}$$

As was done for the isothermal case, the critical speeds in this case are obtained by solving (25).

The function $A_1(\xi)$, occurring in (24), is determined by considering the second order terms. However, if we continue the analysis for the general atmospheres it soon becomes intractable. We shall therefore restrict ourselves to the special case of a linearly increasing temperature distribution in the equilibrium state.

3. Linearly increasing temperature profile

Here we consider the special case of a linearly increasing temperature distribution in the equilibrium state, given by

$$T_e = 1 + \alpha Y, \tag{26}$$

where α is a positive constant. Then (26) together with (5) and the perfect gas relation gives

$$P_e = \lambda (1 + \alpha Y)^{-\mu}, \tag{27}$$

and

$$R_e = (1 + \alpha Y)^{-(1+\mu)} \tag{28}$$

where the definition $\mu = 1/\alpha$ has been introduced for convenience in writing. For Y fixed, the limit $\alpha \to 0$ gives the isothermal distributions of P_e and R_e of *I*.

The equation for P_1, corresponding to (10), may now be obtained as

$$\frac{\partial^3 P_1}{\partial \xi \, \partial Y^2} + \frac{(1 + 2a)}{(1 + aY)} \frac{\partial^2 P_1}{\partial \xi \, \partial Y} + \left[\frac{L(a\gamma + \gamma - 1)}{\gamma \bar{U}_0^2 (1 + aY)}\right] \frac{\partial P_1}{\partial \xi} = 0. \tag{29}$$

The boundary condition $\bar{V}_1(\xi, 0) = 0$ becomes

$$\gamma \frac{\partial^2 P_1}{\partial \xi \, \partial Y} + \frac{\partial P_1}{\partial \xi} = 0 \text{ at } Y = 0. \tag{30}$$

In this case, η is given by

$$\eta = \mu [(\gamma - 1) + \gamma\mu] [(1 + aY)^{1+\mu}], \tag{31}$$

and the condition (20) becomes

$$\int_0^\infty \frac{dY}{\bar{U}_0^2(Y)} < \infty. \tag{32}$$

Under the condition (32), we have two linearly independent solutions, $f(Y)$ and $g(Y)$, of the differential equation

$$\frac{d^2 \bar{F}}{d Y^2} + \frac{(1 + 2a)}{(1 + aY)} \frac{d\bar{F}}{dY} \left[\frac{L(a\gamma + \gamma - 1)}{\gamma \bar{U}_0^2 (1 + aY)}\right] \bar{F} = 0, \tag{33}$$

such that

$$f(Y) = 0 [(1 + aY)^{-(1+\mu)}], \; f'(Y) = 0 [(1 + aY)^{-(2+\mu)}],$$
$$g(Y) = 0(1), \; g'(Y) = 0 [(1 + aY)^{-1}] \text{ as } Y \to \infty. \tag{34}$$

The solution for $P_1(\xi, Y)$ becomes

$$P_1(\xi, Y) = A_1(\xi) f(Y) - (1 + aY)^{-\mu}, \tag{35}$$

where we have used the equilibrium condition on P_1.

Substituting for $P_1(\xi, Y)$ from (35) in (30) we get the following equation for the critical speeds

$$f(0) + \gamma f'(0) = 0. \tag{36}$$

The first order quantities are found to be

$$P_1(\xi, Y) = A_1(\xi) f(Y) - (1 + aY)^{-\mu}, \tag{37}$$

$$R_1(\xi, Y) = - A_1(\xi) f'(Y)/L, \tag{38}$$

$$\bar{V}_1(\xi, Y) = - \frac{A_1'(\xi) [f(Y) + \gamma(1 + aY) f'(Y)] (1 + aY)^{1+\mu} \bar{U}_0}{[(\gamma - 1) L + \gamma L a]}, \tag{39}$$

and

$$\bar{U}_1(\xi, Y) = A_1(\xi) \left[\frac{\bar{U}_0 [f(Y) + \gamma(1 + aY) f'(Y)] (1 + aY)^{1+\mu}}{(\gamma - 1) L + \gamma L a} - \frac{(1 + aY)^{1+\mu} f(Y)}{\bar{U}_0}\right]. \tag{40}$$

The equations governing the second order terms are the same as in *I* with $G_1 - G_4$ defined in the appendix.

The boundary condition $\bar{V}(\xi, 0) = 0$ gives

$$\bar{V}_2(\xi, 0) = 0. \tag{41}$$

The equation for P_2 may be derived by eliminating $\bar{U}_2$, $\bar{V}_2$ and R_2:

$$\frac{\partial^3 P_2}{\partial Y^2 \partial \xi} + \frac{(1 + 2a)}{(1 + aY)} \frac{\partial^2 P_2}{\partial Y \partial \xi} + \left[\frac{L(\gamma - 1 + \gamma a)}{\gamma (1 + aY) \bar{U}_0^2 (Y)}\right] \frac{\partial P_2}{\partial \xi} = G_5, \tag{42}$$

here

$$G_5 = -\left[\frac{L(\gamma - 1 + \gamma a)}{\gamma (1 + aY)}\right]\left(\frac{U_0 G_1 - G_2}{\bar{U}_0^2}\right) + \left[\frac{\gamma a + \gamma - 1}{\gamma (1 + aY)}\right] \frac{\partial G_3}{\partial \xi}$$
$$+ \frac{1}{\gamma (1 + aY)} \frac{\partial}{\partial Y}\left[\gamma (1 + aY) \frac{\partial G_3}{\partial \xi} + \frac{(1 + aY)^{1+\mu} G_4}{\bar{U}_0 (Y)}\right]. \tag{43}$$

The general solution of (42) is given by

$$\frac{\partial P_2}{\partial \xi} = B_1(\xi) f(Y) + B_2(\xi) g(Y) + g(Y) \int_0^Y \frac{G_5(\xi, Y') f(Y')}{W(f, g)(Y')} dY'$$
$$- f(Y) \int_0^Y \frac{G_5(Y') g(Y')}{W(f, g)(Y')} dY', \tag{44}$$

where $B_1(\xi)$ and $B_2(\xi)$ are to be found from the boundary conditions. We shall now analyse the various terms in (44) for large Y. The expression for G_5 given in (43) involves G_1, G_2, G_3 and G_4. As Y becomes large, one may check that the term which contributes most to G_5 is proportional to $\bar{U}_0^2 f(Y)/(1 + aY)$ and this, from (34), is of the order of $\bar{U}_0^2 (1 + aY)^{-(2+\mu)}$ as $Y \to \infty$. On the other hand $W(f, g)(Y)$, the Wronskian of the two linearly independent solutions $f(Y)$ and $g(Y)$ of (33), is easily seen to be proportional to $(1 + aY)^{-(2+\mu)}$. Thus, as $Y \to \infty$, the leading terms in $G_5(\xi, Y) f(Y)/[W(f, g)(Y)]$ and $G_5(\xi, Y) g(Y)/[W(f,g)(Y)]$ would be proportional to $\bar{U}_0^2(Y)(1 + aY)^{-(1+\mu)}$ and $\bar{U}_0^2(Y)$ respectively. Assuming now that $\bar{U}_0(Y)$ is such that

$$\int_0^\infty \bar{U}_0^2(Y')(1 + aY')^{-(1+\mu)} dY' < \infty, \tag{45}$$

we find $\partial P_2/\partial \xi = 0(1)$ as $Y \to \infty$ unless we choose

$$B_2(\xi) = -\int_0^\infty \frac{G_5(\xi, Y') f(Y')}{W(f, g)(Y')} dY'. \tag{46}$$

We take this to be the case and obtain

$$\frac{\partial P_2}{\partial \xi} = B_1(\xi) f(Y) + g(Y) \int_\infty^Y \frac{G_5(\xi, Y') g(Y')}{W(f, g)(Y')} dY'$$
$$- f(Y) \int_\infty^Y \frac{G_5(\xi, Y') g(Y')}{W(f, g)(Y')} dY'. \tag{47}$$

It is easy to see from the asymptotic behaviour of the integrands in (47), given earlier, that $\partial P_2/\partial\xi$ is still not of the same order as $\partial P_1/\partial\xi$ as $Y \to \infty$. For, if we take a linear ambient velocity profile $u_0(y) = \beta y$, so that $\bar{U}_0(Y) = (\beta HY/c - 1)$, and $0 < \alpha < \frac{1}{2}$, we see that the conditions (32) and (45) are satisfied, and

$$\left|\frac{\partial P_2}{\partial\xi}\Big/\frac{\partial P_1}{\partial\xi}\right| = 0\,(Y^3) \text{ as } Y \to \infty. \tag{48}$$

Consequently, if we take P_2 as given by the integral of (47) with respect to ξ, the perturbation expansion for P will be valid only for a limited range of Y; the range will depend on the ambient velocity profile. In the above example of a linearly increasing ambient velocity profile the perturbation expansion for P will be valid, to first order, for all Y which satisfy $Y^3 = 0\,(1/\epsilon)$. In our subsequent analysis we shall restrict our discussion to distances Y, determined by the ambient wind, for which the perturbation expansions remain uniformly valid.

The second order term, $\bar{V}_2$, is then obtained as

$$\begin{aligned}\frac{\bar{V}_2}{\bar{U}_0} = &-\frac{(1+\alpha Y)^{1+\mu}}{[(\gamma-1)+\gamma\alpha]\,L}\Bigg[\{f(Y)+\gamma(1+\alpha Y)f'(Y)\}\Bigg\{B_1(\xi)\\ &-\int_0^Y \frac{G_5\, g(Y')}{W(f,g)(Y')}\,dY'\Bigg\} + \{g(Y)+\gamma(1+\alpha Y)\,g'(Y)\}\\ &\times \int_\infty^Y \frac{G_5 f(Y')}{W(f,g)(Y')}\,dY'\Bigg] + \frac{(1+\alpha Y)^{2+2\mu}\,G_4(\xi, Y)}{L\,[(\gamma-1)+\gamma\,\alpha]\bar{U}_0}\\ &+\frac{\gamma\,(1+\alpha Y)^{2+\mu}}{L\,[\gamma-1)+\gamma\alpha]}\,\frac{\partial G_3}{\partial\xi}.\end{aligned} \tag{49}$$

Applying the condition (41) we get

$$\begin{aligned}&\left[\frac{g(0)+\gamma g'(0)}{d}\right]\int_0^\infty [(1+\alpha Y)^{2+\mu} f(Y)\,G_5]\,dY\\ &+\frac{G_4(\xi,0)}{\bar{U}_0(0)} + \gamma\,\frac{\partial G_3}{\partial\xi}(\xi,0) = 0,\end{aligned} \tag{50}$$

where we have made use of the condition (36) and where $d = W(f,g)(0)$. We now specify the functions $f(Y)$ and $g(Y)$ by requiring $d = 1$ and $g(0) + \gamma g'(0) = 1$. Equation (50), after some lengthy calculations, reduces to

$$m_2\frac{\partial A_1}{\partial\xi} + m_1 A_1\frac{\partial A_1}{\partial\xi} + m_0\frac{\partial^3 A_1}{\partial\xi^3} = 0, \tag{51a}$$

where

$$m_0 = -\int_0^\infty \frac{(1+\alpha Y)^{1+\mu}\,\bar{U}_0^2(Y)\,[f+\gamma\alpha(1+\alpha Y)f']^2}{\gamma L\,[(\gamma-1)+\gamma\alpha]}\,dY, \tag{51b}$$

$$m_1 = \frac{(\gamma - 1)}{\gamma L} f^2(0) + \int_0^\infty (1 + \alpha Y)^{2+\mu} f(Y) [B_1 f^2 + B_2 ff' + B_3 (f')^2]$$
$$dY, \tag{51c}$$

$$m_2 = \frac{[(\gamma - 1) + \gamma\alpha]}{\gamma} \int_0^\infty \frac{(1 + \alpha Y)^{1+\mu} f^2(Y)}{\bar{U}_0^2(Y)} dY, \tag{51d}$$

$$B_1 = -\frac{(3 + \alpha)(1 + \alpha Y)^{\mu-1}}{\gamma \bar{U}_0^2} + \frac{4\bar{U}_0'(1 + \alpha Y)^{\mu}}{\gamma \bar{U}_0^3}$$
$$- \frac{L[\gamma\alpha + (\gamma - 1)](1 + \alpha Y)^{\mu}}{\gamma \bar{U}_0^4} + \frac{(1 + \alpha)(1 + \alpha Y)^{\mu-2}}{[(\gamma - 1) + \alpha\gamma] L},$$

$$B_2 = \frac{2(1 + \alpha Y)^{\mu-1}}{L[(\gamma - 1) + \gamma\alpha]} + \frac{4\bar{U}_0'(1 + \alpha Y)^{1+\mu}}{\bar{U}_0^3} + \frac{(2\alpha - 4/\gamma)(1 + \alpha Y)^{\mu}}{\bar{U}_0^2},$$

and

$$B_3 = \left\{\frac{1 + 2\alpha^2\gamma + \alpha\gamma - 2\alpha}{L[(\gamma - 1) + \gamma\alpha]}\right\}(1 + \alpha Y)^{\mu} - \frac{(1 + \alpha Y)^{1+\mu}}{\bar{U}_0^2}.$$

It is easy to see that, as $\alpha \to 0$, (51a) reduces to (43a) with its co-efficients given by (43b)–(43d) of I. The latter corresponds to the equation for solitary and cnodial waves for an isothermal atmosphere. As in the isothermal case, the solitary and cnoidal wave solutions of (51a) may be written out. We may also show that the speed of the solitary waves, c, which is close to a critical speed c^*, is such that $|c| > |c^*|$.

Equations (37)–(40) give the first order solution with $A_1(\xi)$ satisfying (51a). The quantities, P_1, R_1, $\bar{U}_1$ and $\bar{V}_1$, are of the same order, for all Y, as their zeroth order counterparts. However, the second order solutions grow in relation to the first order quantities as $Y \to \infty$, the growth depending on the ambient velocity profile. Consequently, the solutions given by (37)–(40) would remain uniformly valid only in a limited range of Y provided by the ambient shear flow.

4. An example

We now apply the results of § 3 to the special case

$$u_0(y) = \beta y, \ \beta > 0. \tag{52}$$

The critical speeds can be obtained by solving the differential equation

$$\frac{d^2\bar{F}}{dY^2} + \frac{(1 + 2\alpha)}{(1 + \alpha Y)}\frac{d\bar{F}}{dY} + \frac{Ri_m}{(1 + \alpha Y)[Y - c^*/(\beta H)]^2}\bar{F} = 0, \tag{53}$$

together with the conditions

$$\bar{F} = 0\,[1/(1 + \alpha Y)^{1+\mu}] \text{ as } Y \to \infty, \tag{54a}$$

$$\bar{F}(0) + \gamma \bar{F}'(0) = 0, \tag{54b}$$

where

$$Ri_m = (\alpha\gamma + \gamma - 1)\, g/\gamma H\beta^2, \tag{55}$$

is the maximum Richardson number. Equation (53) follows from (33) and (52). We preclude critical levels, thus requiring that the two singularities of (53) do not lie in the flow field.

Equation (53) can be transformed to the hypergeometric equation

$$Z(1 - Z)\frac{d^2 G}{dZ^2} + [2k - \{1 + (k + 1 + \mu) + k\} Z]\frac{dG}{dZ}$$
$$- k(k + 1 + \mu)\, G = 0, \tag{56}$$

by introducing the new variables

$$G(Z) = (Y - c^*/(\beta H))^{-k}\, \bar{F}(Y), \tag{57}$$

$$Z = -\alpha(Y - c^*/(\beta H))/\delta \tag{58}$$

where

$$k = 1/2 \pm (1/4 - Ri_m/\delta)^{1/2}, \tag{59}$$

and $\delta = 1 + \alpha c^*/(\beta H)$.

The boundary conditions (54a, b) become

$$G = 0\,(1/[(1 + \alpha Y)^{1+\mu}\;(Y - c^*/(\beta H))^k]) \text{ as } Y \to \infty, \tag{60}$$

and

$$\left[\left(\frac{1-\delta}{\delta}\right) + \gamma k\right] G\left(\frac{\delta - 1}{\delta}\right) + \gamma\left(\frac{\delta - 1}{\delta}\right) G'\left(\frac{\delta - 1}{\delta}\right) = 0. \tag{61}$$

The solution of (56) which has the asymptotic behaviour given by (60) is

$$G(Z) = (Y - c^*/(\beta H))^{-(1+\mu+k)}\, F(a', b';\; c';\; Z^{-1}), \tag{62}$$

where

$$a' = 3/2 + \mu + ik_2,\; b' = 3/2 + \mu - ik_2,\; c' = 2 + \mu,$$

and $k_2 = (Ri_m/\delta - 1/4)^{1/2}$.

The nature of the eigenvalues is obtained by substituting $G(Z)$ from (62) into (61) to get

$$\bar{G}(\delta) = [(1 - \delta)^2 - \gamma\alpha(1 + 1/\alpha)(1 - \delta)]\, F(a', b';\; c';\; \delta/(\delta - 1))$$
$$+ \gamma\alpha\delta F'(a', b';\; c';\; \delta/(\delta - 1)) = 0, \tag{63}$$

where the prime over F, the hypergeometric function, denotes the derivative with respect to the last argument. Since we require c^* to be negative, to avoid critical levels, we seek the roots δ of (63) which lie between 1 and $-\infty$. We also observe that the function $\bar{G}(\delta)$, given in (63), is analytic in δ for δ belonging to the range $(-\infty, 1)$ and, consequently, the limit points of the zeros of $\bar{G}(\delta)$, if any, could only be $\delta = -\infty$ or $\delta = 1$. Three distinct cases arise depending on whether the maximum Richardson number, Ri_m, is less than, equal to, or greater than 1/4. The nature of the eigenvalues can be predicted by knowing the asymptotic behaviour

of the hypergeometric function for the above three cases and the analysis is similar to the one presented in paper I, the details may be found in [4]. The number of roots of $\bar{G}(\delta)$ in (63) is finite when Ri_m is less than or equal to 1/4, while there is an infinity of them in the neighbourhood of $\delta = 1$ for Ri_m greater than 1/4.

While these results are obtained from the eigenvalue problem, our theory, as in the isothermal case, does not include solitary waves travelling near critical speeds c_n^* for which n is arbitrarily large. This is because for large n, c_n^*, which is negative, is very close to zero and the origin $Y = 0$ comes close to the critical level $Y = c^*/(\beta H)$, which lies below the ground. It is easily verified that, for large n, R_1 and $\bar{U}_1$ at the origin, given by (38) and (40) become large. Hence the spectrum of possible critical speeds, as obtained from the eigenvalue problem, has to be restricted so that the critical level is not in the close vicinity of the flow field.

5. Conclusions

In the present paper we have extended the results of I to more general atmospheres. We have considered an atmosphere in which, initially, the pressure and density were strictly decreasing with Y and the temperature was a non-decreasing function of Y. With the help of a simple transformation, a growth condition on the ambient wind profile was obtained which ensured that the first order terms are of the same order as the zeroth order ones. The problem of finding the critical speeds in this case also reduced to solving an eigenvalue problem for a second order ordinary differential equation. Further analysis in this general case, however, was found to become too involved and therefore, for simplicity, a linearly increasing temperature profile was assumed. This corresponds to assuming the atmosphere to be a pure thermosphere. Results similar to those stated in the isothermal case were obtained. Also, for a linear wind profile, the first order equations posed an eigenvalue problem for the Gauss's hypergeometric equation whose eigenvalues give the possible critical speeds of the flow. In this case the Richardson number varied with height and decreased to zero as Y increased to infinity. We have shown that if the maximum Richardson number, Ri_m, is greater than 1/4, then the number of possible critical speeds is infinite. On the other hand, if $Ri_m \leqslant 1/4$, only a finite number of them were found to exist. This change in the behaviour of the number of critical speeds with the maximum Richardson number can be explained on the basis of the behaviour of the solutions near the critical level, as for the isothermal case. Also, our comments made in the isothermal case, on the extent of the validity of our solution and the finiteness of the spectrum of the critical speeds, if we wish to stay away from the critical levels, remain valid in this case too. Therefore, generally, the limited range of validity of the solution as well as the need to keep the critical level out of the domain of the solution both inhibit the cardinality of the solitary waves so that even for $Ri_m > 1/4$ we shall have only a small number of possible solitary waves in the atmosphere.

Acknowledgement

This work was in part carried out under the University Grants Commission project 'Wave Phenomena' No. 23-884/78 CSRIF-Code No. 001/Maths/78.

Appendix

In this appendix we give the G_i's for the case of linear temperature profile. Writing $\eta = (1 + \alpha Y)$, and using the notations as in I, we have,

$$G_1 = \left[\frac{\eta^\mu}{\gamma L \bar{U}_0} f^2 + \frac{\eta^\mu}{L^2(\gamma - 1 + \gamma\alpha)} (\alpha \bar{U}_0 + \eta \bar{U}_0') ff' + \frac{\eta^{1+\mu}}{L^2(\gamma - 1 + \gamma\alpha)} (\gamma\eta\bar{U}_0' + (\gamma - 1 + 2\gamma\alpha) \bar{U}_0) f'^2\right] A_1 \frac{\partial A_1}{\partial \xi},$$

$$G_2 = \Bigg[\Bigg\{ -\left(\frac{\bar{U}_0'}{L(\gamma - 1 + \gamma\alpha)} - \frac{1}{\bar{U}_0}\right)^2 \eta^{1+\mu} + \frac{\bar{U}_0}{L(\gamma - 1 + \gamma\alpha)}\left(\frac{(\alpha + 1)\eta^\mu \bar{U}_0'}{L(\gamma - 1 + \gamma\alpha)} + \frac{\eta^{1+\mu}\bar{U}_0''}{L(\gamma - 1 + \gamma\alpha)} - \frac{(\alpha + 1)\eta^\mu}{\bar{U}_0}\right)\Bigg\} f^2 + \Bigg\{\frac{\bar{U}_0}{L(\gamma - 1 + \gamma\alpha)}\left(\frac{2\gamma\eta^{2+\mu}\bar{U}_0''}{L(\gamma - 1 + \gamma\alpha)} + \frac{(\gamma + 1 + \gamma\alpha)\eta^{1+\mu}\bar{U}_0'}{L(\gamma - 1 + \gamma\alpha)} - \frac{2\gamma(\alpha + 1)\eta^{1+\mu}}{\bar{U}_0}\right) - 2\left(\frac{\bar{U}_0'}{L(\gamma - 1 + \gamma\alpha)} - \frac{1}{\bar{U}_0}\right)\left(\frac{\gamma\eta^{2+\mu}\bar{U}_0'}{L(\gamma - 1 + \gamma\alpha)}\right)\Bigg\} ff' + \Bigg\{\frac{\gamma\bar{U}_0}{L(\gamma - 1 + \gamma\alpha)}\left(\frac{\eta^{2+\mu}\bar{U}_0'}{L(\gamma - 1 + \gamma\alpha)} + \frac{\gamma\eta^{3+\mu}\bar{U}_0''}{L(\gamma - 1 + \gamma\alpha)} - \frac{\eta^{2+\mu}}{\bar{U}_0} - \frac{\gamma^2\eta^{3+\mu}}{L^2(\gamma - 1 + \gamma\alpha)^2}\right)\Bigg\} f'^2\Bigg] A_1 \frac{\partial A_1}{\partial \xi},$$

$$G_3 = -\frac{A_1 f'}{L} + \frac{\bar{U}_0^2(f + \gamma\eta f')}{L(\gamma - 1 + \gamma\alpha)} \frac{\partial^2 A_1}{\partial \xi^2},$$

$$G_4 = -\left(\frac{f\bar{U}_0}{L\eta^{1+\mu}}\right)\frac{\partial A_1}{\partial \xi} + \frac{1}{L(\gamma - 1 + \gamma\alpha)}\Bigg[\left\{\frac{\gamma(\alpha + 1)\bar{U}_0}{\eta} - \bar{U}_0'\right\} f^2 + \{(1 + \gamma - \gamma\alpha)\bar{U}_0 - 2\gamma\eta\bar{U}_0'\} ff' + \{2\gamma\eta\bar{U}_0 - \gamma^2\eta^2\bar{U}_0' - \gamma^2(1 + 2\alpha)\eta\bar{U}_0\} f'^2\Bigg] A_1 \frac{\partial A_1}{\partial \xi},$$

and

$$G_5 = \left[\frac{(\gamma - 1 + \gamma\alpha)}{\gamma\eta\bar{U}_0^2} f\right]\frac{\partial A_1}{\partial \xi} + (c_1 f^2 + c_2 ff' + c_3 f'^2) A_1 \frac{\partial A_1}{\partial \xi} + \frac{1}{L(\gamma - 1 + \gamma\alpha)}\left[\left\{\frac{(\gamma - 1 + 2\gamma\alpha)\bar{U}_0^2}{\gamma\eta} + 2\alpha\bar{U}_0\bar{U}_0' - 1\right\} f + \{\gamma\alpha + 2\gamma\eta\bar{U}_0\bar{U}_0'\} f'\right]\frac{\partial^3 A_1}{\partial \xi^3},$$

where c_1, c_2 and c_3 are given by

$$c_1 = \frac{(a+1)\,\eta^{\mu-2}}{L(\gamma-1+\gamma a)} - \frac{(a+3)\,\eta^{\mu-1}}{\gamma \bar{U}_0^2} + \frac{4\eta^{\mu}\,\bar{U}_0'}{\gamma \bar{U}_0^3} - \frac{L(\gamma-1+\gamma a)}{\gamma \bar{U}_0^4}\,\eta^{\mu},$$

$$c_2 = \frac{2\eta^{\mu-1}}{L(\gamma-1+\gamma a)} + \frac{(2a-4/\gamma)\,\eta^{\mu}}{\bar{U}_0^2} + \frac{4\eta^{1+\mu}\,\bar{U}_0'}{\bar{U}_0^3},$$

and

$$c_3 = \frac{(1+2a^2\gamma + a\gamma - 2a)\eta^{\mu}}{((\gamma-1)+\gamma a)\,L} - \frac{\eta^{1+\mu}}{\bar{U}_0^2}.$$

References

[1] Abramowitz M and Stegun I A (eds) 1965 *Handbook of mathematical functions* (New York : Dover Publications)

[2] Hille E 1969 *Lectures on ordinary differential equations* (Massachusetts : Addison-Wesley) pp. 428–432

[3] Sachdev P L and Seshadri V S 1979 *Phys. Fluids* **22** 60

[4] Seshadri V S 1977 Some analytic approach to nonlinear wave propagation in non-uniform media Ph.D. Thesis, Indian Institute of Science, Bangalore

Proc. Indian Acad. Sci. (Math. Sci.), Vol. 91, Number 1, January 1982, pp. 53–57.

On minimizing the duration of transportation

SATYA PRAKASH
Department of Mathematics, B.I.T.S., Pilani, 333 031, India

MS received 6 October 1980 ; revised 14 December 1981

Abstract. The problem of minimizing the duration of transportation has been studied. The problem has been reduced to a goal programming-type problem which readily lends itself to solution by the standard transportation method. This approach to the solution of the problem is very much different from all other existing ones.

Keywords. Optimisation ; goal programming ; transportation.

1. Introduction

The problem of minimizing the duration of transportation has been dealt by many workers —Hammer [3], [4], Garfinkel and Rao [2], Szwarc [9], Bhatia *et al* [1], Ramakrishnan [6], Sharma and Swarup [8], Seshan and Tikekar [7]. The present paper also deals with this problem. The problem has been reduced to a goal programming-type problem which readily lends itself to solution by the standard transportation method. This approach to the solution of the problem is very much different from all other existing ones. A numerical example has been given to clarify the procedure to solve the problem toward the end.

2. Formulation of the problem

Suppose that there are m origins and n destinations. Given amounts of a uniform product are available at the origins and specified amounts of the product are required at the destinations. It is possible to transport the product from any origin to any destination. Further, the transportation starts simultaneously and the time of transportation from any origin to any destination does not depend on the amount of the product transported. Let t_{ij} be the time of transportation of the product from origin i to destination j, a_i the units of the product available at origin i, b_j the units of the product required at destination j, and x_{ij} the number of units of the product transported from origin i to destination j. It is required to determine the minimum duration routing from the origins to the destinations subject

to the above mentioned constraints. The mathematical formulation of this problem is as follows. Find $x_{ij} \geqslant 0$ $(i = 1, \ldots, m;\ j = 1, \ldots, n)$ which minimize

$$z = \max \{t_{ij} : x_{ij} > 0\ (i = 1, \ldots, m;\ j = 1, \ldots, n)\}, \tag{1}$$

subject to the constraints

$$\sum_{j=1}^{n} x_{ij} = a_i\ (i = 1, \ldots, m), \tag{2}$$

$$\sum_{i=1}^{m} x_{ij} = b_j\ (j = 1, \ldots, n). \tag{3}$$

3. Solution procedure

The objective function of the problem formulated above is not linear and so the problem is not linear. Below we outline a procedure to obtain the solution of this problem. We reduce this problem to a goal programming-type problem discussed by Hughes and Grawiog [5]. To do this we partition the set $\{t_{ij} :\ i = 1, \ldots, m;\ j = 1, \ldots, n\}$ into subsets L_k $(k = 1, \ldots, q)$ in the following way. Each of the subsets L_k consists of the t_{ij} having the same numerical value. L_1 consists of the t_{ij} having the greatest value, L_2 consists of the t_{ij} having the next greatest value, and so on. Finally L_q consists of the t_{ij} having the lowest value. After this, we associate a cost M_k with each of the x_{ij} corresponding to the t_{ij} belonging to L_k $(k = 1, \ldots, q)$. The M_k are all positive numbers.

With all this done, we are now in a position to formulate an alternative problem whose optimal basic feasible solution would yield the desired solution. The alternative problem seeks to determine $x_{ij} \geqslant 0$ $(i = 1, \ldots, m;\ j = 1, \ldots, n)$ which minimize

$$z = C\Big(\sum_{k=1}^{q} M_k \sum_{L_k} x_{ij}\Big), \tag{4}$$

subject to the constraints (2) and (3). Here $\sum_{L_k} x_{ij}$ refers to the summation over the x_{ij} corresponding to the t_{ij} belonging to L_k and C is a function such that

$$C\Big(\sum_{k=1}^{q} M_k \sum_{L_k} x_{ij}\Big) = M_s \sum_{L_s} x_{ij} \text{ where } s = \min \{k : \sum_{L_k} x_{ij} \neq 0\ (k = 1, \ldots, q)\} \tag{5}$$

The function C in (5) qualitatively treats each M_k as very much larger than M_{k+1}.

It is to be noted that the given problem is equivalent to the alternative problem because minimizing the objective function of both the problems determines $x_{ij} \geqslant 0$ $(i = 1, \ldots, m;\ j = 1, \ldots, n)$ for which the duration of transportation of the product from the origins to the destinations would be minimum. Hereinafter we shall refer to the alternative problem by the name 'equivalent problem'. The equivalent problem is a goal programming-type problem which readily lends itself to solution by the standard transportation method.

4. A numerical example

Now we shall apply the above procedure to obtain the solution of a numerical problem which is obtained by taking $m = 4$, $n = 6$ and assigning numerical values to all other quantities in the problem formulated above in §2. The tableau representation of this numerical problem is shown in table 1.

Table. 1. Tableau representation of the numerical problem.

	P_1	P_2	P_3	P_4	P_5	P_6	a_i
R_1	25	30	20	40	45	37	37
R_2	30	25	20	30	40	20	22
R_3	40	20	40	35	45	22	32
R_4	25	24	50	27	30	25	14
b_j	15	20	15	25	20	10	

In this table, row i of the table is denoted by R_i and its column j is denoted by P_j. The cells in the table correspond to the variables x_{ij} $(i = 1, \ldots, 4;\ j = 1, \ldots, 6)$. The left top corner of all the cells (i, j) depicts the units of time of transportation of the product from origin $i\,(i = 1, \ldots, 4)$ to destination $j\,(j = 1, \ldots, 6)$. The marginal column depicts the units of the product available at the origins and the marginal row depicts the units of the product required at the destinations. To obtain the units of the product available at origin i, sum the x_{ij} associated with the cells (i, j) across the row. To obtain the units of the product required at destination j, sum the x_{ij} associated with the cells (i, j) across the column. And the objective function for this numerical problem is given by

$$z = \max\ \{t_{ij} : x_{ij} > 0\ (i = 1, \ldots, 4;\ j = 1, \ldots, 6)\}. \tag{6}$$

For the numerical problem, we find $q = 11$.

The subsets forming the partition of the set $\{t_{ij} : i = 1, \ldots, 4;\ j = 1, \ldots, 6\}$ for the numerical problem are as follows:

$$L_1 = \{t_{43}\},\ L_2 = \{t_{15},\ t_{35}\},\ L_3 = \{t_{14},\ t_{25},\ t_{31},\ t_{33}\},\ L_4 = \{t_{16}\},\ L_5 = \{t_{34}\},$$
$$L_6 = \{t_{12},\ t_{21},\ t_{24},\ t_{45}\},\ L_7 = \{t_{44}\},\ L_8 = \{t_{11},\ t_{22},\ t_{41},\ t_{46}\},\ L_9 = \{t_{42}\},$$
$$L_{10} = \{t_{36}\},\ L_{11} = \{t_{13},\ t_{23},\ t_{26},\ t_{32}\}.$$

The t_{ij} belonging to $L_1, L_2, L_3, L_4, L_5, L_6, L_7, L_8, L_9, L_{10}, L_{11}$ have the values 50, 45, 40, 37, 35, 30, 27, 25, 24, 22, 20 respectively.

After this we associate a cost M_k with each of the x_{ij} corresponding to the t_{ij} belonging to the set L_k $(k = 1, \ldots, 11)$.

The tableau representation of the equivalent problem associated with the numerical problem is shown in table 2.

Table 2. Representation of the equivalent problem associated with the numerical problem.

	P_1	P_2	P_3	P_4	P_5	P_6	a_i
R_1	M_8 (15)	M_6	M_{11} (15)	M_3	M_2 (7)	M_4	37
R_2	M_6	M_8	M_{11}	M_6 (11)	M_3 (11)	M_{11}	22
R_3	M_3	M_{11} (20)	M_3	M_5	M_2 (2)	M_{10} (10)	32
R_4	M_8	M_9	M_1	M_7 (14)	M_6	M_8	14
b_j	15	20	15	25	20	10	

It is to be noted that the left top corner of the cells (i, j) in this table no longer depicts the time of transportation but it depicts the cost. For this equivalent problem, the objective function which we seek to minimize is

$$z = C(M_1 x_{43} + M_2 (x_{15} + x_{35}) + M_3 (x_{14} + x_{25} + x_{31} + x_{33}) + M_4 x_{16} + \\ + M_5 x_{34} + M_6 (x_{12} + x_{21} + x_{24} + x_{45}) + M_7 x_{44} + M_8 (x_{11} + x_{22} + \\ x_{41} + x_{46}) + M_9 x_{42} + M_{10} x_{36} + M_{11} (x_{13} + x_{23} + x_{26} + x_{32})). \quad (7)$$

Here C is the function defined by (5) with $q = 11$.

The equivalent problem is a goal programming-type problem which is amenable to solution by the standard transportation method. Applying the standard transportation method, we obtain the optimal basic feasible solution of the equivalent problem. An initial basic feasible solution for this equivalent problem is obtained by applying the column-minima method and the values of the basic variables of this initial solution are entered inside circles in table 2. Skipping the routine intermediate steps, the final tableau giving the optimal basic feasible solution is sown in table 3. In this table, all the $(z_{ij} - c_{ij})$ corresponding to the nonbasic cells are nonpositive after the application of the C function as in (5) indicating that the basic feasible solution is optimal. The values of the basic variables of the optimal basic feasible solution are entered inside circles in the table. Further, we see that the greatest cost M_3 is associated with $\bar{x}_{25}$ among all the non-zero basic variables. It may be noted that M_3 is the cost associated with the x_{ij} corresponding to the t_{ij} belonging to the set $L_3 = \{t_{14}, t_{25}, t_{31}, t_{33}\}$ each of whose elements

Table 3. The optimal basic feasible solution

	P_1	P_2	P_3	P_4	P_5	P_6	a_i
R_1	M_8; (15)	M_6; (7)	M_{11}; (15)	M_3; $-M_3+M_5+M_6-M_{11}$	M_2; $-M_2+M_3+M_5-M_{11}$	M_4; $-M_4+M_6+M_{10}-M_{11}$	37
R_2	M_6; $-M_5-M_6+M_8+M_{11}$	M_8; $-M_5+M_6-M_8+M_{11}$	M_{11}; $-M_5+M_{11}$	M_6; (16)	M_3; (6)	M_{11}; $-M_5+M_6+M_{10}-M_{11}$	22
R_3	M_3; $-M_3-M_6+M_8+M_{11}$	M_{11}; (13)	M_3; $-M_3-M_6+2M_{11}$	M_5; (9)	M_2; $-M_2+M_3+M_5-M_6$	M_{10}; (10)	32
R_4	M_8; $-M_3-M_5+M_6+M_{11}$	M_9; $-M_3-M_5+2M_6-M_9+M_{11}$	M_1; $-M_1-M_3-M_5+M_6+2M_{11}$	M_7; $-M_3+2M_6-M_7$	M_6; (14)	M_8; $-M_3-M_5+2M_6-M_8$	14
b_j	15	20	15	25	20	10	

has the value 40. And t_{25} is the units of time of transportation of the product from origin 2 to destination 5. So the duration of transportation of the product for the desired solution is $t_{25} = 40$ units.

Acknowledgements

The author thanks the UGC for providing financial assistance and to the BITS at Pilani for providing facilities for carrying out this work. Thanks are also due to the referee for suggesting certain improvements.

References

[1] Bhatia H L, Swarup K and Puri M C 1977 *Indian J. Pure Appl. Math.* **8** 920

[2] Garfinkel R S and Rao M R 1971 *Nav. Res. Logist Q.* **18** 465

[3] Hammer P L 1969 *Nav. Res. Logist Q.* **16** 345

[4] Hammer P L 1971 *Nav. Res. Logist Q.* **18** 487

[5] Hughes A J and Grawiog D E 1973 *Linear Programming : An emphasis on Decision making* (USA : Addison-Wesley) pp. 300–312

[6] Ramakrishnan C S 1977 *Opsearch* **14** 207

[7] Seshan C R and Tikekar V G 1980 *Proc. Indian Acad. Sci. (Math. Sci.)* **89** 101

[8] Sharma J K and Swarup K 1977 *Proc. Indian Acad. Sci. (Math. Sci.)* **86** 513

[9] Szwarc W 1971 *Nav. Res. Logist Q.* **18** 473

Proc. Indian Acad. Sci. (Math. Sci.), Vol. 91, Number 1, January 1982, pp. 59–63.

Weyl's theorem and thin spectra

SHANTI PRASANNA

Department of Mathematics, University School of Sciences, Gujarat University, Ahmedabad 380 009, India

MS received 24 April 1981

Abstract. Using the notion of thin sets we prove a theorem of Weyl type for the Wolf essential spectrum of $T \in \beta(H)$. *Further we show that Weyl's theorem holds for a restriction convexoid operator and consequently modify some results of Berberian. Finally we show that Weyl's theorem holds for a paranormal operator and that a polynomially compact paranormal operator is a compact perturbation of a diagnoal normal operator. A structure theorem for polynomially compact paranormal operators is also given.

Keywords. Hilbert space operators; spectrum essential spectra; Weyl's theorem; testriction convexoid operators; paranormal operators.

1. Weyl's theorem

Weyl showed that if T is a self-adjoint operator then $\sigma(T) \setminus \omega(T) = \pi_{00}(T)$. Coburn [8] extended this result to seminormal and Toeplitz operators. Berberian [3] showed that if T is restriction convexoid and is reduced by all its finite dimensional eigenspaces then Weyl's theorem holds for T. Since the validity of Weyl's theorem plays a vital role in determining the normality of certain classes of operators it is desirable to know whether Weyl's theorem holds for operators for which one may not know whether the eigenspaces are reducing. In this paper using the notion of thin sets we prove a theorem of Weyl type for the Wolf essential spectrum.

Definition 1.1: $T \in \beta(H)$ is said to satisfy α_0 if every isolated point of $\sigma(T)$ is a pole of order one of the resolvent $R_\lambda(T)$.

Definition 1.2: Thin sets [5]. If σ is a nonempty closed subset of the extended complex plane we denote by $A(\sigma)$ the algebra of σ-analytic functions. σ is said to be thin if $A(\sigma) = C(\sigma)$.

**Notations* We denote by $\beta(H)$ the Banach algebra of bounded linear operators on a complex Hilbert space H. For $T \in \beta(H)$ we denote the spectrum and the point spectrum of T by $\sigma(T)$ and $\pi_0(T)$ respectively. The symbols $\hat{\sigma}(T)$, $\hat{\sigma}_l(T)$ and $\omega(T)$ stand for the Wolf essential spectrum, the left essential spectrum and the Weyl spectrum of T respectively. Weyl's theorem is said to hold for T if $\sigma(T) \setminus \omega(T) = \pi_{00}(T)$ where $\pi_{00}(T)$ is the set of isolated eigenvalues of T with finite geometric multiplicity. For the operational calculus methods used here refer to [13]. Unless otherwise stated H is not assumed to be separable.

Theorem 1.1 : Let H be a separable Hilbert space. If T satisfies α_0 and if $\sigma(T)$ is thin, then

(1) $\sigma(\hat{T}) = \sigma_l(\hat{T}) = \omega(T)$

(2) $\omega(T) = \sigma(\hat{T}) = \sigma(T)\setminus\pi_{00}(T)$

(3) $\omega(T^*) = \sigma(\hat{T^*}) = \sigma(T^*)\setminus\pi_{00}(T^*)$.

Proof : Let $\lambda_0 \in \pi_{00}(T)$. Since T satisfies α_0, λ_0 is a pole of $R_\lambda(T)$ of order 1. Hence by theorem 5·8*A* of [13], $R(T-\lambda_0) = N(B_1)$ where B_1 is the spectral projection associated with λ_0. Thus $R(T-\lambda_0)$ is closed. Also by the same theorem of [13] ascent of $T-\lambda_0$ is 1. Hence $N(T-\lambda_0)^n = N(T-\lambda_0)$ for all n or $\dim N(T-\lambda_0)^n = \dim N(T-\lambda_0)$ is a finite number. Thus $\lim_{n\to\infty} \dim N(T-\lambda_0)^n$ is finite. Applying theorem 4·5 of [12], $T-\lambda_0$ is a Fredholm operator of zero index. Hence

$$\lambda_0 \notin \omega(T) \text{ or } \pi_{00}(T) \subset \sigma(T)\setminus\omega(T) \subset \sigma(T)\setminus\sigma(\hat{T}) \subset \sigma(T)\setminus\sigma_l(\hat{T}). \quad (1)$$

Our next step is to show that $\sigma(T) = \sigma_l(\hat{T}) \cup \pi_{00}(T)$. Since $\sigma_l(\hat{T})$ and $\pi_{00}(T)$ are subsets of $\sigma(T)$, it is sufficient to show that $\sigma(T) \subset \sigma_l(\hat{T}) \cup \pi_{00}(T)$. Since $\sigma(T)$ is thin its interior is empty [5]. Thus every point of $\sigma(T)$ is either an isolated point of $\sigma(T)$ or is a boundary point of $\sigma(T)$.

Case 1 : Let λ_0 be an isolated point of $\sigma(T)$. Since T satisfies α_0, λ_0 is a pole of $R_\lambda(T)$ and consequently is an eigenvalue of T. If $N(T-\lambda_0)$ is finite dimensional then $\lambda_0 \in \pi_{00}(T)$. If not then $\lambda_0 \in \sigma_l(\hat{T})$ [9]. Thus $\lambda_0 \in \sigma_l(\hat{T}) \cup \pi_{00}(T)$.

Case 2 : Let λ_0 be a boundary point of $\sigma(T)$. Since $\partial\sigma(T) \subset \sigma_l(\hat{T}) \cup \pi_{00}(T)$, [9] it follows that $\lambda_0 \in \sigma_l(\hat{T}) \cup \pi_{00}(T)$.

Thus in either cases $\sigma(T) \subset \sigma_l(\hat{T}) \cup \pi_{00}(T)$ or $\sigma(T) = \sigma_l(\hat{T}) \cup \pi_{00}(T)$.

Hence

$$\sigma(T)\setminus\sigma_l(\hat{T}) \subset \pi_{00}(T). \quad (2)$$

From 1 and 2 we have,

$$\sigma(\hat{T}) = \sigma_l(\hat{T}) = \omega(T) \quad \text{and} \quad \sigma(T)\setminus\pi_{00}(T) = \sigma(\hat{T}) = \omega(T).$$

Since $\sigma(T^*)$ is thin so is $\sigma(T^*)$. Suppose λ_0 is an isolated point of $\sigma(T^*)$. Then $\bar{\lambda}_0$ is an isolated point of $\sigma(T)$ and consequently is a pole of order 1 of $R_\lambda(T)$. Hence λ_0 is a pole of order 1 of $R_\lambda(T^*)$. Thus the conditions of the theorem are satisfied for T^* also and (3) holds.

Corollary 1.1 : If $T \in \beta(H)$ satisfies α_0 and if $\sigma(T)$ has planar Lebesque measure zero then Weyl's theorem holds for T.

Proof: Since $\sigma(T)$ has planar measure zero by Hartogs-Rosenthal theorem, $\sigma(T)$ is thin [5] and the result follows from theorem 1.1.

2. Restriction convexoid operators

Following Bouldin [6] we define an asymptotic eigenspace of an operator as follows:

Definition 2.1: Suppose corresponding to $\lambda \in \sigma(T)$ there exists a number $\delta < 1$ such that $|(f, g)| \leqslant \delta$ whenever $f \in N(T-\lambda)$. $\|f\| = 1 = \|g\|$ and g is an eigenvector for some eigenvalue of T distinct from λ. Then we say that $N(T-\lambda)$ is not an asymptotic eigenspace of T. If this condition fails then $N(T-\lambda)$ is called an asymptotic eigenspace.

Bouldin characterized the Browder and Weyl essential spectra based on the above notion and showed.

Theorem A: If T is a restriction spectraloid operator then T has no asymptotic eigenspaces.

Before proceeding further we introduce some notations that will be needed in this and the ensuing section.

If λ_0 is an isolated point of $\sigma(T)$, the dimension of $R(B_1)$ is called the algebraic multiplicity of λ_0.

Define $\hat{\pi}_{00}(T)$ $[\hat{\pi}_{0f}(T)] = \pi_{00}(T)$ $[\pi_{0f}(T)]$ minus Isolated eigenvalues of T with infinite algebraic multiplicity where π_{0f} (T) denotes the eigenvalues of T with finite geometric multiplicity.

Baxley (1972) extended the results of Bouldin and introduced the notions C_{-1} and C_{-2} for an operator. T is said to satisfy: C_{-1} if $\{\lambda_n\} \in \hat{\pi}_{00}(T)$, $\lim\limits_{n\to\infty} \{\lambda_n\} = \lambda \in \hat{\pi}_{0f}(T)$ and if $\{x_n\}$ is a sequence of corresponding normalised eigenvectors then $\{x_n\}$ does not converge and

$$C_{-2}: \text{ if } \lambda \in \pi_{00}(T) \text{ implies } R(T-\lambda) \text{ is closed.}$$

Baxley proved the following theorems:

Theorem B: If each finite dimensional eigenspace of T is not an asymptotic eigenspace then T satisfies C_{-1} and

Theorem C: If T satisfies C_{-1} and C_{-2} then $\sigma(T) \diagdown \omega(T) = \pi_{00}(T)$.

Using these theorems we show below that Weyl's theorem holds for a restriction convexoid operator.

Lemma 2.1: If T is restriction convexoid then T satisfies a_0 and consequently C_{-2}.

Proof: Let λ_0 be an isolated point of $\sigma(T)$. Since $\sigma(T/R(B_1)) = \{\lambda_0\}$ and $T/R(B_1)$ is convexoid, $T/R(B_1) = \lambda_0 I/R(B_1)$. Hence $(T-\lambda_0)B_1 = 0$ or λ_0 is a pole of $R_\lambda(T)$ of order 1. Thus $R(T-\lambda_0)$ is closed by theorem 5·8 A of [13] or T satisfies C_{-2}.

Theorem 2.1 : If T is a restriction convexoid operator then Weyl's theorem holds for T.

Proof : Since T is restriction convexoid, T is restriction spectraloid. Hence by theorems A and B, T satisfies C_{-1}. Also by lemma 2.1, T satisfies C_{-2}. Hence Weyl's theorem holds for T by theorem C.

Berberian showed that if T is (1) restriction convexoid (2) Weyl's theorem holds for T and (3) $\omega(T)$ is finite then T is polynomially compact [4]. In view of theorem 2.1 the condition (2) is redundant and a stronger assertion holds :

We denote by $\sigma(T)'$ the limit points of $\sigma(T)$.

Theorem 2.2 : Let T be a restriction convexoid operator. Then T is polynomially compact if and only if $\omega(T)$ is finite.

Proof : Suppose T is polynomially compact. Let $p(\lambda)$ be the polynomial such that $p(T)$ is compact. Since $p(T)$ is compact, 0 is the only possible limit point of $\sigma(p(\tau))$. Hence by spectral mapping theorem $\sigma(T)'$ is finite. Theorem 3 of [2] now implies that T is normal. Hence $\sigma(\hat{T}) = \omega(T) \subset \{\lambda \mid p(\lambda) = 0\}$ or $\omega(T)$ is finite.

Conversely suppose $\omega(T)$ is finite. Then since Weyl's theorem holds for T and $\sigma(T)' \subset \sigma(T) \setminus \pi_{00}(T) = \omega(T)$, $\sigma(T)'$ is finite. Hence T is normal by theorem 3 of [2]. That T is polynomially compact follows from the finiteness of $\omega(T) = \sigma(\hat{T})$.

Corollary 2.1 : Let T be a restriction convexoid operator. If $\omega(T) = \{0\}$ then T is normal and compact.

3. Paranormal operators

While discussing hyponormal operators and related topics Saito [11] posed the question whether every polynomially compact paranormal operator is normal. We show in this section that though such an operator may fail to be normal it differs from a normal operator only by a compact operator.

Theorem 3.1 : If T is paranormal then Weyl's theorem holds for T.

Proof : Since T is paranormal T satisfies a_0 [10] and consequently C_{-2}. Further T being restriction spectraloid by theorems A and B T satisfies C_{-1} also. Hence by theorem C Weyl's theorem holds for T.

Theorem 3.2 : Let T be a polynomially compact paranormal operator on a separable Hilbert space H. Then $T = N + k$ where N is polynomially compact normal operator and k is compact. Moreover one can write $T = \sum_{i=1}^{m} \oplus (K_i + \lambda_i) + K$ where the K_i's are diagonal normal operators and $\sigma(\hat{T}) = \omega(T) = \{\lambda_1, \lambda_2, \ldots, \lambda_m\}$.

Proof: Since T is paranormal and polynomially compact T satisfies α_0 and $T^*T - TT^*$ is compact [10] Also $\sigma(T)$ has planar Lebesque measure zero. Hence by theorem 1.1, $\sigma(\hat{T}) = \omega(T)$. Thus $\lambda \notin \sigma(\hat{T})$ implies index of $T - \lambda$ is zero. Hence by a theorem of Brown, Douglas and Fillmore [7], $T = N + K$ where N is normal and K is compact. Since $p(T) = p(N)$ + terms involving K, N is polynomially compact. Hence by theorem 3 of [4], $N = \sum_{i=1}^{m} \oplus (K_i + \lambda_i)$ where the K_i's are compact normal operators and $\sigma(\hat{T}) = \sigma(\hat{N}) = \{\lambda_1, \lambda_2, \ldots, \lambda_m\}$.

Acknowledgement

The author would like to thank Dr I H Sheth for his guidance and encouragement throughout this work.

References

[1] Baxley J V 1972 *Proc. Am. Math. Soc.* **34**

[2] Berberian S K 1970 *Math. Ann.* **184** 181–192

[3] Berberian S K 1969 *Mich. Math. J.* **16** 273–279

[4] Berberian S K 1970 *Proc. Am. Math. Soc.* **26** 277–281

[5] Berberian S K 1974 *Lectures in functional analysis and operator theory* (New York: Springer Verlag)

[6] Bouldin R 1971 *Proc. Am. Math. Soc.* **28**

[7] Brown L G, Douglas R G and Fillmore P A 1973 *Unitary equivalence modulo the compact operators and extensions of C* algebras.* Proceedings of a Conference on Operator Theory (New York: Springer Verlag) **345**

[8] Coburn L A 1966 *Mich. Math. J.* **13** 285–288

[9] Fillmore P A, Stampfli J G and Williams J P 1972 *Acta Sci. Math.* **33** 179–192

[10] Prasanna S and Sheth I H To appear in *J. Ind. Math. Soc.*

[11] Saito T 1972 *Hyponormal operators and related topics* (New York: Springer Verlag) **247** 534–664

[12] Schechter M 1972 *Principles of functional analysis* (New York: Academic Press)

[13] Taylor A E 1963 *Introduction to functional analysis* (New York: John Wiley)

[14] Weyl H 1909 *Rend. Circ. Mat. Palermo* **27** 373–392

Proc. Indian Acad. Sci. (Math. Sci.), Vol. 91, Number 1, January 1982, pp. 65–71.

On the normality of the rings of Schubert varieties

C HUNEKE and V LAKSHMIBAI
Department of Mathematics, The University of Michigan, Ann Arbor, Michigan 48109, USA

MS received 26 May 1981

Abstract. We prove that the cone over a Schubert variety in G/P (P being a maximal parabolic subgroup of classical type) is normal by exhibiting a 2-regular sequence in $R(w)$ (the homogeneous coordinate ring of the Schubert variety $X(w)$ in G/P under the canonical projective embedding $G/P \hookrightarrow (P(H^\circ G/P, L))$, L being the ample generator of (PicG/P), which vanishes on the singular locus of $X(w)$. We also prove the surjectivity of $H^\circ(G/Q, L)$ $H^\circ(X(w), L)$, where Q is a classical parabolic subgroup (not necessarily maximal) of G and L is an ample line bundle on G/Q.

Keywords. Schubert varieties; parabolic subgroups; standard monomials; normality.

1. Introduction

Let G be a semi-simple, simply-connected algebraic group defined over an arbitrary field k. Let T be a maximal torus, B, a Borel subgroup containing T. Let Q be a parabolic subgroup, containing B; further, let $Q = \bigcap_{i=1}^{r} P_i$, for some maximal parabolic subgroups $P_1, P_2, \ldots, P_r$ of G. Let W be the Weyl group of G relative to T and W_Q, that of Q. For $w \in W/W_Q$, let $X(w) = \overline{BwQ}$ (mod Q) with the canonical reduced scheme structure be the Schubert subvariety of G/Q associated to w. Let $L = L_1^{a_1} \otimes L_2^{a_2} \otimes \ldots \otimes L_r^{a_r}$ be a line bundle on G/Q where L_i is the ample generator of Pic (G/P_i), $1 \leqslant i \leqslant r$; further let $L \underset{\neq}{>} 0$ (by which we mean $a_i \underset{\neq}{>} 0$, $1 \leqslant i \leqslant r$) so that L is very ample on G/Q. For the canonical projective embedding $G/Q \hookrightarrow P(H^\circ(G/Q, L))$, let $R(w)$ denote the homogeneous coordinate ring of $X(w)$. One would like to ask the question "Is the ring $R(w)$ normal?" The answer to this question is known in the following cases.

Case 1.1

$G = SL_n$ and Q is a maximal parabolic subgroup (*i.e.*, for Schubert varieties in the Grassmann variety). The proof in this case follows from the results of [4], [13] or [14] wherein $R(w)$ is proved to be Cohen–Macaulay (in [4], the author

also proves the normality of $R(w)$) and this together with Chevalley's result (cf. [1]) that Schubert varieties are non-singular in codimension 1 gives the normality of $R(w)$ in this case. For the special case of $X(w)$ being the Grassmann variety itself, the normality of $R(w)$ is also proved in [5].

Case 1.2

Q is a maximal parabolic subgroup of classical type. This case is proved in [2] (refer [11] or [12], for the definition of a maximal parabolic subgroup of classical type). In the special case of Q being a maximal parabolic subgroup of quasi-minuscule type, the above result is also proved in [10].

Case 1.3

G is a classical group and $Q = B$. The proof in this case follows from the results of [8], for the class of Kempf varieties. For $G = SL(n)$ and for the class of Kempf varieties, this is also proved in [6]. Again, for G, a classical group and $X(w) = G/B$, this is also proved in [9].

The proof of normality of $R(w)$ (in case 1.2 above) in the spirit of [2] consists in first proving that $R(w)$ is Cohen–Macaulay and then deducing normality using Chevalley's result mentioned above. Now, in the first part of this paper, we prove the normality (for the case 1.2 above, which of course, includes case 1.1) by exhibiting a 2-regular sequence in $R(w)$ which vanishes on the singular locus of $X(w)$. (This was motivated by Hochster).

In the second part of this paper, we set aright a slightly non-trivial gap in [11] (G/P-IV), which we briefly describe below.

Let $Q = P$, a maximal parabolic subgroup of classical type and let L be the ample generator of Pic (G/P). Let $P_{\tau,\phi}$ be the elements of $H^\circ(X(w), L)$ as defined in [11] (p. 324, theorem 5.10) (where $w \geq \tau$). A monomial $p_{\tau_1,\phi_1} \cdot p_{\tau_2,\phi_2} \cdot \ldots\, p_{\tau_m,\phi_m}$ is said to be standard on $X(w)$, if $w \geq \tau_1 \geq \phi_1 \geq \tau \geq \phi_2$ and so on (cf. [11]). Then we have that the standard monomials on $X(w)$ give a basis for $R(w)$ (cf. [11], p. 332 prop. 7.3). Then using Lemma 8.2 of [11] and Grothendieck's criterion (cf. [7], p. 415), one concludes that $\bigoplus_{m\geq 0} H^\circ(G/P, L^m) \to \bigoplus_{m\geq 0} H^\circ(X(w), L^m)$ is surjective. (This surjectivity is concluded in [11], by claiming (*cf.* [11], proof of Prop. 8.3), that Lemma 8.2 of [11] implies the normality of $R(w)$; but this claim is not true as was pointed out by Hochster). Now the surjectivity implies that $R(w)$ is in fact $= \bigoplus_{m\geq 0} H^\circ(X(w), L^m)$ and hence in particular, the space $H^\circ(X(w), L^m)$ has a basis consisting of standard monomials of degree m on $X(w)$.

Now let Q be a non-maximal parabolic subgroup of classical type, say $Q = \bigcap_{i=1}^{r} P_i$, where each P_i is a maximal parabolic subgroup of classical type. Let $S(w) = \bigoplus_{L\geq 0} H^\circ(X(w), L)$ where $X(w)$, as above is a Schubert variety in G/Q and $L = \bigotimes_{i=1}^{r} L_i^{a_i}$, L_i being the ample generator of Pic (G/P_i). Let $A(w)$ be the sub-algebra of $S(w)$ generated by the elements $p_{\theta,\sigma}$ of $S(w)$ (refer [11], p. 341, Defn. 10.5, for the definition of $p_{\theta,\sigma}$). Then (as in the maximal parabolic case) we have that $A(w)$ has a basis consisting of standard monomials (which are defined in terms of $p_{\theta,\sigma}$'s refer [11], defn. 10.5). To conclude that $S(w)$ has a basis consisting of standard

monomials, one may either prove a similar surjectivity as above or otherwise. Now in [11], the required surjectivity is concluded using an analogous version of *Lemma* 8.2 of [11], for the case of a non-maximal parabolic subgroup. Unlike the maximal parabolic case (as discussed above, where the Grothendieck's criterion together with *Lemma* 8.2 of [11] gave the required surjectivity) one does not obtain the required surjectivity, in the non-maximal parabolic subgroup case. In the latter half of this paper, we remove this gap by showing that $A(w)$ is in fact $=$ $S(w)$, which then yields the required surjectivity. In [15], the authors conclude the required surjectivity as follows. Firstly, one has the surjectivity $H^\circ(G_{\mathbf{Z}}/Q_{\mathbf{Z}}$ $L_{\mathbf{Z}}) \to H^\circ(X_{\mathbf{Z}}(w), L_{\mathbf{Z}})$ over Z (as a consequence of the standard monomial theory, *cf.* [11]). And now this surjectivity implies Demazure's conjecture, (*cf* [3], p. 83) the proof being the same as in [11], p. 337, remark 9.6. Now, Demazure's conjecture implies that $H^i(X(w), L) = 0, i \geq 1$ and as a consequence of this, one obtains a basis for $H^\circ(X(w), L)$ in terms of standard monomials and hence, one also obtains that $H^\circ(G/Q, L) \to H^\circ(X(w), L)$ is surjective. The normality of $S(w)$ (in the non-maximal parabolic subgroup case) is still an open problem.

2. Normality of $R(w)$

Let G be a semi-simple simply-connected algebraic group over a field k, T a maximal torus B, a Borel subgroup $\supseteq T$, W, the Weyl group, P a maximal parabolic subgroup of classical type (*cf.* [11], [12]), L the ample generator of Pic (G/P), and W_p the Weyl group of P. For $w \in W/W_p$, let $X(w) = \overline{BwP}$ (mod P), together with the canonical reduced scheme structure, be the Schubert subvariety of G/P, associated to w. For the canonical projective embedding

$$G/P \hookrightarrow P(H^\circ(G/P, L))$$

let $R(w)$ be the homogeneous coordinate ring of $X(w)$. Let $p_{\tau,\phi}$ (with $w \geq \tau$) be the elements of $H^\circ(X(w), L)$ as defined in §5 of [11]. Call a monomial $p_{\tau_1,\phi_1} \cdot p_{\tau_2,\phi_2} \cdots p_{\tau_r,\phi_r}$ to be *standard on* $X(w)$ if $w \geq \tau_1 \geq \phi_1 \geq \tau_2 \geq \phi_2 \geq \cdots$ then we have that the monomials standard on $X(w)$ give a basis for $R(w)$ (*cf.* [11], prop. 7.3). Also, as already remarked in the introduction, the *Lemma* 8.2 of [11] implies that $H^\circ(G/P, L^m) \to H^\circ(X(w), L^m)$ is surjective. In particular we obtain $H^\circ(X(w)), L^m) = (R(w))_m$ and hence it has a basis consisting of monomials of degree "m" standard on $X(w)$. Let us denote $R = R(w)$, for simplicity.

Proposition 2.1. Let $f = \Sigma p_{w_i}$, where the summation runs over all w_i such that $X(w_i)$ is a Schubert subvariety of $X(w)$ of co-dimension 1. Then f is not a zero divisor in $R/(p_w)$.

Proof: Let F be a non-zero element in $R/(p_w)$. Writing F as a sum of monomials standard on $X(w)$, we have

$$F = \sum_{\lambda} a_\lambda\, p_{w,\lambda}, g_\lambda + \sum_{\tau \lneq w} b_{\tau,\mu} p_{\tau,\mu} h_{\tau,\mu}$$

$$(\mathrm{I} \quad + \qquad \mathrm{II})$$

Case 1. Suppose all expressions in II are zero, i.e., $b_{\tau,\mu} = 0$ for all (τ, μ). Then $F = \sum_{\lambda} a p_{w,\lambda} g_{\lambda}$. Now fix a λ such that $a_{\lambda} \neq 0$. Then, if $w_i \not\geq \lambda$, then $p_{w_i} p_{w,\lambda} \in (p_w)$, in view of the strong lexicographic order condition in the straightening relations for $p_{w_i} p_{w,\lambda}$ (cf. [2], definition 1.2 and theorem 4.1) and if $w_i \geq \lambda$, then $p_{w,w_i} p_{w_i,\lambda}$ appears with coefficient ± 1 in the expression for $p_{w,\lambda} p_{w_i}$ as a sum of standard bionomials (*cf.* [2], propositions 4.2 and 4.4). Hence, the expression for F as a sum of standard monomials involves non-zero terms like $a_{\lambda} p_{w,w_i} p_{w_i,\lambda} g_{\lambda}$ (since for any λ, there exists at least one i with $w_i \geq \lambda$) from which the proposition follows.

Case 2. There exists at least one $b_{\tau,\mu} \neq 0$. Fix such a $b_{\tau,\mu}$. Then there exists at least one w_i such that $w_i \geqslant \tau$. And now $p_{w_i} p_{\tau,\mu} h_{\tau,\mu}$ is standard. Now we make the *claim* : The term $b_{\tau,\mu} p_{w_i} p_{\tau,\mu} \cdot h_{\lambda,\mu}$ does not get cancelled, in the expression for fF as a sum of standard monomials. *Proof of the claim* : Firstly, observe that $p_{w_i} p_{\tau,\mu} h_{\tau,\mu}$ cannot appear in the expression for $f \sum_{\lambda} p_{w,\lambda} g_{\lambda}$ as a sum of standard monomials (in view of the strong lexicographic order condition (*cf.* [2], defn. 1.2 and theorem 4.1).

Secondly, consider $p_{w_k} \cdot p_{\tau',\mu'} \cdot h_{\tau'\mu'}$. Either this remains standard (namely when $w_k \geqslant \tau'$) ; in the alternate case, every monomial, occurring in the expression for $p_{w_k} \cdot p_{\tau',\mu'} \cdot h_{\tau',\mu'}$ as a sum of standard monomials, starts with $p_{w,\theta}$ for some $\theta \leqslant w$, in view of the lexicographic order condition (*cf.* [2], defn. 1.2 and theorem 4.1).

Now the claim follows from the above two observations and it is obvious that the claim implies the proposition.

Proposition 2.2. The elements p_w and $f(= \sum p_{w_i})$ of $H^{\circ}(X(w), L)$ vanish on the singular locus of $X(w)$.

Proof: In view of Chevalley's result (*cf.* [1]) that Schubert varieties are non-singular in codim 1, we would be done, for instance, if we show that both fw and f vanish on Schubert subvarieties of codim 2 in $X(w)$. But this is immediate in view of the fact that for $\theta, \phi \in W/W_P$, $p_{\theta} \mid_{X(\phi)} \neq 0$ if and only if $\phi \geqslant \theta$ (*cf.* [11]-p. 324, theorem 5.10).

Proposition 2.3. The ring R is normal.

Proof: Let p be any prime ideal of R. If $(p_w, f) \subseteq p$, then R_p has depth $\geqslant 2$ (in view of proposition 2.1). If $(p_w, f) \not\subseteq p$, then $p \not\supseteq I$(singular locus of $X(w)$) (in view of prop. 2.2). Hence R_p is regular ; in particular R_p is normal and hence R_p has depth at least two if the codimension of p is at least two. Thus if the codimension of p is at least two, the depth R_p is also at least two. This together with the fact that R is non-singular in codimension one implies that R is normal (*cf.* [16]).

3. Surjectivity of $H^\circ(G/Q, L) \to H^\circ(X(w), L)$

Let Q be a parabolic subgroup of classical type (*cf.* [11]), i.e., $Q = \bigcap_{i=1}^{r} P_i$ where P_i's are maximal parabolic subgroups of classical type. Let L_i be the ample generator of Pic (G/P_i), $1 \leqslant i \leqslant r$. Let $L = L_1^{a_1} \otimes L_2^{a_2} \otimes .. \otimes L_n^{a_r}$, with $a_i \geqslant 0, 1 \leqslant i \leqslant r$. We shall refer to such an L being positive. Let $S(w) = \bigoplus_{L \geq 0} H^\circ(X(w), L)$. Let $A(w)$ be the sub algebra of $S(w)$, which as a subspace is spanned by $p_{\phi,\theta}$ (*cf.* [11], defn. 10.5). Then we have that $A(w)$ has a basis consisting of standard monomials (*cf.* [11]). For simplicity, let us denote $S = S(w)$ and $A = A(w)$.

Lemma **3.1.** Let R be an integral domain and let $x, y, z, \in R$. Then

(a) (x, yz) is a regular 2-sequence if and only if (x, y) and (x, z) are

(b) (xy, z) is a regular 2-sequence if and only if (x, z) and (y, z) are.

Proof: Is easy.

Lemma **3.2.** (Generalization of Lemma 3.1). Let $\{x_i\}$ $1 \leqslant i \leqslant m$, $\{y_j\}$ $1 \leqslant j \leqslant r$ be elements of an integral domain R. Then $(x_1 x_2 \ldots x_m y_1 y_2 \ldots y_r)$ is a regular 2-sequence if and only if (x_i, y_j)'s are $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant r$.

Proposition 3.3. For $w \in W/W_Q$, let $X(w_i)$ be the projection of $X(w)$ under $G/Q \to G/P_i$, $1 \leqslant i \leqslant r$ (where $Q = \bigcap_{i=1}^{r} P_i$). Then in $A (= A(w))$ $(p_{w_1}^{a_1} \otimes p_{w_2}^{a_2} \otimes \ldots \otimes p_{w_r}^{a_r}, \; p_{Id}^{b_1} \otimes p_{Id}^{b_2} \otimes \ldots \otimes p_{Id}^{b_r})$ is a regular 2-sequence where a_i and b_i are $\geqslant 0$, $1 \leqslant i \leqslant r$.

Proof: In view of Lemma 3.2 enough to check that $1 \otimes \ldots \otimes p_{w_i}^{a_i} \otimes 1 \otimes \ldots \otimes 1$, $1 \otimes 1 \otimes \ldots \otimes p_{Id}^{b_j} \otimes \ldots \otimes 1)$ is a regular 2-sequence. Since the construction of the basis for A by means of standard monomials (*cf.* [11]) does not depend on the ordering of the maximal parabolics $(P_1, \ldots, P_r)$ we may suppose $i = 1$ and $j = r$ (in other words, we have a basis for A consisting of standard monomials corresponding to the ordering $(P_i, \ldots, P_j)$ of the maximal parabolics $(P_1, \ldots, P_r)$.

Now, if possible, let F, C, in R be such that $F\, p_{Id}^{b_r} = p_{w_1}^{a_1} G$. Writing $F = \sum_k F_k$, $G = \sum_t G_t$ (sum of standard monomials on $X(w)$) we have, $\Sigma F_k \cdot p_{Id}^{b_r} = \Sigma p_{w_1}^{a_1} G_t$ wherein both LHS and RHS are standard on $X(w)$. Hence, linear independence of standard monomials (*cf.* [11]) implies that for each k, there exists a t such that $F_k \cdot p_{Id}^{b_r} = p_{w_1}^{a_1} G_t$. And this implies that $p_{w_1}^{a_1} \mid F_k$, $\forall k$.

This completes the proof of proposition 3.3.

Proposition 3.4. Quotient field of A = quotient field of S (A and S as defined in the beginning of this section).

Proof: Given an s-tuple $(a_{i_1}, a_{i_2}, \ldots, a_{i_s})$ where $0 \leqslant s \leqslant r$ and $a_{ij} \underset{\neq}{>} 0$, $1 \leqslant j \leqslant s$ if $L = L_{i_1}^{a_{i_1}} \otimes L_{i_2}^{a_{i_2}} \otimes \ldots \otimes L_{i_s}^{a_{i_s}}$, $Q_s = \cap P_j$ $j = i_1, i_2, \ldots, i_s$ and $X(\tau)$ is the

projection of $X(w)$ under $G/Q \to G/Q_s$, considering the projective embeddings $X(\tau) \hookrightarrow P(H^\circ(G/Q_s, L))$ for various choices of the tuples $(a_{i_1}, a_{i_2} \ldots, a_{i_s})$, $0 \leqslant s \leqslant r$, we conclude (in view of Serre's Theorem for projective varieties) $A_{b_{i1}, b_{i2}, \ldots, b_{is}} = S_{b_{i1}, b_{i2}, \ldots, b_{is}}$ if G.C.D. $(b_{i_1}, \ldots, b_{i_s}) \gg 0$. Now let $v \in S_{m_{i1}, \ldots, m_{is}}$, $N = m_{i_1} m_{i_2} \ldots m_{i_s}$ and $q \gg 0$. Choose $0 \neq u \in A_{N^q m_{i_1}}, N^q m_{i_2}, \ldots, N^q_{m_s} m_{i_s} (= S_{N^q m_{i_1}, N^q m_{i_2}, \ldots, N^q m_{i_s}})$.
Then $vu \in A$ (since G.C.D. $(N^q m_{i_1} + m_{i_1}, N^q m_{i_2} + m_{i_2}, \ldots, N^q m_{i_s} + m_{i_s})$ has $N^q + 1$ as a divisor. And now, $v = vu/u$ belongs to the quotient field of A. Also $S_0 = A_0$.

This completes the proof of proposition 3.4.

Proposition 3.5. Let R be an integral domain, with quotient field k. Let $t \in k$ be such that $(u, v)\, t \subseteq R$, where (u, v) is a regular 2-sequence in R. Then $t \in R$.

Proof: Let $t = a/b$, $ut = s_1$, $vt = s_2$. Then we have $ua = s_1 b$; $va = s_2 b$; hence $uas_2 = vas_1$ which gives $us_2 = vs_1$. This in turn implies that $s_1 = r.u$ for some $r \in R$ (since (u, v)) is a regular 2 sequence in R). Hence we obtain $ut\ (= s_1) = ru$ which gives $t = r$, as required.

Proposition 3.6: We have $A = S$ (A and S being as in proposition 3.4).

Proof: Let $s \in S$, so that $s \in k$, the quotient field of A (which is also the quotient field of S, *cf.* proposition 3.4 above). Further, let $s \in S_{m_{i_1}, \ldots, m_{i_k}}$. May choose

$$u = p_{w_{i1}}^{b_{i1}} \otimes p_{w_{i2}}^{b_{i2}} \otimes \ldots \otimes p_{w_{ik}}^{b_{ik}}, \text{ and } v = p_{Id}^{b_{i1}} \otimes p_{Id}^{b_{i2}} \otimes \ldots \otimes p_{Id}^{b_{ik}},$$

where $b_{i_t} = N^q m_{i_t}$, $l \leqslant t \leqslant k$ (N and q as in the proof of proposition 3·4) so that su and sv are both in A and we are through in view of propositions 2·3 and 3.5 above).

Proposition 3.7. Let $L = L_1^{a_1} \otimes \ldots \otimes L_r^{a_r}, a_i \geqslant 0$. Then the canonical map

$$H^\circ(G/Q, L) \to H^\circ(X(w), L)$$

is surjective.

This is an easy consequence of proposition 3.6. In particular, we have the following.

Proposition 3.8. Let G be classical (so that any parabolic subgroup of G is of classical type (cf. [11])) and L a positive line bundle on G/B. Then for $w \in W$, the canonical map

$$H^\circ(G/B, L) \to H^\circ(X(w), L)$$

is surjective.

Remark 3.9. As a consequence of proposition 3.8, we get a proof of Demazure's conjecture (*cf.* [3], p. 83) for the case of a classical group.

References

[1] Chevalley C C 1958 *Circa* (manuscript non publié)

[2] Deconcini C and Lakshmibai V Arithmetic Cohen-Macaulayness and arithemetic normality of Schubert varieties. To appear. AJM

[3] Demazure M 1974 *Ann. Sec. Ec Norm. Sup.* t7 53–88

[4] Hochster M 1973 *J. Algebra* **25** 40–57

[5] Igusa J I 1954 *Proc. Natl. Acad. Sci., U.S.A.* **40** 309–313

[6] Kempf G R 1976 *Am. J. Math.* **98** 325–331

[7] Kleiman S L and Landolfi J 1973 *Compos. Math.* **23** 407–434

[8] Lakshmibai V 1976 *J. Indian Math. Soc.* **40** 299–349

[9] Lakshmibai V, Musili C and Seshadri C S 1974 *Ann. Sec. Ec. Norm. Sup.* **t7** 89–138

[10] Lakshmibai V, Musili C and Seshadri C S 1978 *Proc. Indian Acad. Sci.* **A87** 93–177

[11] Lakshmibai V, Musili C and Seshadri C S 1979 *Proc. Indian Acad. Sci.* **A88** 279–362

[12] Lakshmibai V and Seshadri C S 1978 *Proc. Indian Acad. Sci.* **A87** 1–54

[13] Laksov D 1972 *Acta Math.* **129** 1–9

[14] Musili C 1972 *J. Indian Math. Soc.* **36** 143–171

[15] Musili C and Seshadri C S Standard Monomial Theory (a survey) Lecture notes in Mathematics (New York: Springer Verlag) No. 867 pp. 441–476

[16] Serre J P 1965 *Algèbre Locale Multiplicités, Lecture Notes in Math.* (New York: Springer-Verlag) Vol. 11

Proc. Indian Acad. Sci. (Math. Sci.), Vol. 91, Number 2, July 1982, pp. 73–122.

The fundamental group-scheme*

MADHAV V NORI
School of Mathematics, Tata Institute of Fundamental Research, Colaba, Bombay 400 005, India

MS received 2 November 1981

1. Introduction

Let X be a compact Riemann surface, $\tilde{X}$ a finite Galois unramified covering, with Galois group G. Let V be a vector space with G-action. The diagonal action of G on $\tilde{X} \times V$ is free and the quotient is a vector bundle W on X. It was shown by A Weil that there are two polynomials f and g with non-negative integer coefficients with $f \neq g$ and $f(V)$ isomorphic to $g(V)$. Isomorphism classes of vector bundles on X form a semi-ring with respect to direct sums and tensor products, so the expressions $f(V)$ and $g(V)$ make sense as vector bundles on X.

A vector bundle W on X satisfying this property is called *finite*. We prove the converse : a finite vector bundle W arises from a representation of the Galois group for a suitable unramified covering $\tilde{X} \to X$.

It is easy to see that a line bundle L is finite if and only if L is a point of finite order in the Jacobian of X. For such a line bundle, the function field of $\tilde{X}$ is just a simple Kummer extension of the function field of X. Thus our theorem for line bundles simply asserts that the characters of the etale fundamental group of X (which is the profinite completion of the topological fundamental group of X are in one-to-one correspondence with line bundles of finite order on X. This is, of course, a well-known fact, and a very useful one because the structure of the abelian group of all such line bundles is determined very easily by the topology of the Jacobian. Whereas it is not clear how to go about determining the finite bundles from the variety of stable bundles on X ; consequently our theorem has met with no utility.

If X is a complete connected reduced scheme over a field k, finite vector bundles still make sense. An essentially finite bundle is just a sub-quotient of W, remaining in the semi-stable category. If G is a finite group-scheme and P is a principal G-bundle on X, the representations of G give rise to essentially finite bundles on X and in fact all essentially finite bundles are obtained in this manner. In characteristic zero, finite = essentially finite.

This is the content of Chapter I.

* Ph.D. thesis submitted to University of Bombay.

While proving this, we show that there is an affine group-scheme $\pi(X, \chi_0)$ which is an inverse limit of finite group-schemes, a principal $\pi(X, \chi_0)$-bundle P on X, a base-point $*$ of P sitting above χ_0 which is a k-rational point of X with the following universal property : given a principal G-bundle Q on X with G being a finite group-scheme and V a base-point of Q above χ_0, there is a unique pair (f, ρ) such that

(i) $\rho : \pi(X, \chi_0) \to G$ is a homomorphism. (ii) $f : P \to Q$ interwines the actions of $\pi(X, \chi_0)$ and G, and (iii) $f(*) = V$.

Naturally we call $\pi(X, \chi_0)$ the fundamental group-scheme of X at χ_0.

This leads us to the questions : when does such a $(P, \pi(X, \chi_0), *)$ exist with the above universal property ? In characteristic zero, there is no problem at all : this is just the etale fundamental group.

In Chapter II, we show that $\pi(X, \chi_0)$ exists if X is connected and reduced (the completeness is not necessary). That some conditions on X are necessary was suggested by Milne who showed (in our language) that $\pi(X, \chi_0)_{ab}$ exists if and only if all members of $\Gamma(X, \mathbf{O}_X)$ integral over k belong to k (in particular, $\Gamma(X, \mathbf{O}_X)$ has no nilpotents). It seems unlikely that this $\pi(X, \chi_0)$ has the decent properties enjoyed by the usual fundamental group, e.g.,

A†. If $X \to S$ is a smooth proper morphism and S is not equi-characteristic then the fundamental group-schemes of the fibres certainly do not vary in a flat manner and this destroys some of the interest in this concept.

B. If $P \to X$ is a principal G-bundle on X with G a finite group-scheme, then P may not have a fundamental group-scheme at all !

On the positive side, we have :

A. A proper smooth morphism with connected fibres induces a surjection of fundamental group-schemes.

B. $\pi(X, \chi_0)$ is a birational invariant for smooth complete varieties.

C. It remains unaffected by the removal of a closed subset of codimension $\geqslant 2$ if the ambient space is regular.

D. $\pi(X, \chi_0)$ is trivial for normal rational varieties.

E. It remains invariant under base-change by separable extensions.

The existence of $\pi(X, \chi_0)$ and the proofs of the above statements are deal with in Chapter II.

In Chapter III we show that all the results of Chapter I are valid for parabolic bundles (which are a slight modification of the parabolic bundles defined and used by C S Seshadri). Thus we show that for a smooth projective connected curve X and a finite set $S \to X$, the representations of $\pi(X-S, \chi_0)$ are in one-to-one correspondence with essentially finite parabolic bundles on $X-S$.

In the final chapter we consider principal bundles on X with nilpotent structure-groups. In this context there is again a nilpotent affine group-scheme

† This was shown by M. Artin.

$U(X, \chi_0)$ and a principal bundle P with this structure group and a base-point having the obvious universal property. In characteristic zero, $U(X, \chi_0)$ is determined completely by its Lie algebra. In positive characteristic, however, $U(X, \chi_0)$ is an inverse limit of finite group-schemes and therefore a quotient of $\pi(X, \chi_0)$. We show that

A. $U(X, \chi_0)_{ab} = \varprojlim_{G \to \operatorname{Pic} X} \hat{G}$ where the G run through all local group-schemes embedded in Pic X (assuming that Pic X exists).

B. If X is an abelian variety, $U(X, 0)$ is abelian.

C. If X is an elliptic curve, $\pi(X, 0) \to \varprojlim_{G \to \hat{X}} \hat{G}$ is an isomorphism, where the G run through all finite sub-group-schemes of X.

Equivalently, if P is a principal G-bundle on X with a base-point v above the zero of X and G is a finite group-scheme, then there is a homomorphism $\rho : X_n \to G$ for some n and a commutative diagram :

$$\begin{tikzcd} X \arrow[rr, "f"] \arrow[dr, "n"'] & & P \arrow[dl] \\ & X & \end{tikzcd}$$

such that f intertwines the X_n-action on X and the action of G on P.

We fail to prove however that this holds for abelian varieties.

D. Invariance under arbitrary field extensions.

The remaining part of the chapter is devoted to a preliminary study of $U(X, \chi_0)$ for curves X in positive characteristic. We find that the $U(X, \chi_0)$ are determined by "non-commutative formal groups" which are defined there. A classification of such objects presents an interesting problem. We then compute $U(X, \chi_0)$ for rational curves with rather simple singularities, and also prove an old result of Safarevich.

The appendix gives an easy proof for the results about Tannaka Categories stated in Chapter I, § 1.

Literature referred to is mentioned at the end of each chapter.

Chapter I is a reproduction of "Representations of the Fundamental Group" which appeared in *Compositio Math.*, 1976, Vol. 33. It has been included here for the sake of completeness. Several conventions introduced in Chapter I have been adhered to throughout.

---

* A new proof of the theorem on Tannaka categories has also appeared in Springer Verlag Lecture Notes 900.

PART I

CHAPTER 1

2. On the Representations of the Fundamental Group

2.1. *Tannaka categories*

Let G be an affine group-scheme defined over a field k, R its coordinate ring, and G-mod the category of finite-dimensional left representations of G. Let k-mod be the category of finite-dimensional k-vector spaces, and $T_k : G\text{-mod} \to k\text{-mod}$ the forgetful functor. Let $\hat{\otimes}$ ($\otimes$ resp.) denote the usual tensor product functor on G-mod (k-mod resp.). Let L_0 be the trivial representation of G.

Putting $(G\text{-mod}, \hat{\otimes}, T_k, L_0) = (\mathcal{C}, \hat{\otimes}, T, L_0)$, we note that the following statements are true :

$\mathcal{C}1$: $\mathcal{C}$ is an abelian k-category (existence of direct sums of finite object of $\mathcal{C}$ included).

$\mathcal{C}2$: Obj $\mathcal{C}$ is a set.

$\mathcal{C}3$: $T : \mathcal{C} \to k$-mod is a k-additive faithful exact functor.

$\mathcal{C}4$: $\hat{\otimes} : \mathcal{C} \times \mathcal{C} \to \mathcal{C}$ is a functor which is k-linear in each variable, and $T \circ \hat{\otimes} = \otimes \circ (T \times T)$.

$\mathcal{C}5$: $\hat{\otimes}$ is associative, preserving T, in the following sense : Let $H: \hat{\otimes} \circ \ (I_{\mathcal{C}} \times \hat{\otimes}) \to \hat{\otimes} \circ (\hat{\otimes} \times I_{\mathcal{C}})$ be the equivalence of functors that give the associativity of $\hat{\otimes}$. For objects V_1, V_2, V_3 of $\mathcal{C}$, $T\,(H(V_1, V_2, V_3))$ gives an isomorphism of $TV_1 \otimes (TV_2 \otimes TV_3)$ with $(TV_1 \otimes TV_2) \otimes TV_3$. We ask that this isomorphism coincides with the usual one that gives the associativity of the tensor product for vector spaces.

$\mathcal{C}6$: $\hat{\otimes}$ is commutative, preserving T, in the above sense.

$\mathcal{C}7$: There is an object L_0 of $\mathcal{C}$, and an isomorphism $\varphi : k \to TL_0$, such that L_0 is an identity object of $\hat{\otimes}$, preserving T.

$\mathcal{C}8$: For every object L of $\mathcal{C}$ such that TL has dimension equal to one, there is an object L^{-1} such that $L \hat{\otimes} L^{-1}$ is isomorphic to L_0.

Any $(\mathcal{C}, \hat{\otimes}, T, L_0)$ shall be called a Tannaka category.

Definition : Let $\mathcal{C}$ be any category where $\mathcal{C}1$ and $\mathcal{C}2$ hold. Let S be a subset of Obj $\mathcal{C}$. Then

$$\bar{S} = \{W \in \text{Obj}\,\mathcal{C} : \exists\, P_i \in S,\ 1 \leqslant i \leqslant t, \text{ and } V_1, V_2 \in \text{Obj}\,\mathcal{C} \text{ such that}$$

$$V_1 \subseteq V_2 \subseteq \bigoplus_{i=1}^{t} P_i, \text{ and } W \text{ is isomorphic to } V_2/V_1\}.$$

By $\mathcal{C}(S)$, we mean the full subcategory of $\mathcal{C}$ with Obj $\mathcal{C}(S) = \bar{S}$. Note that $\mathcal{C}(S)$ is also an abelian category. Finally, S will be said to generate $\mathcal{C}$ if obj $\mathcal{C} = \bar{S}$. The following theorems are due to Saavedra (see Theorem 1 of Saavedra [4]).

Theorem (1.1) : Any Tannaka category is the category of finite-dimensional left representations of an affine group-scheme G, and this sets up a bijective correspondence between affine group-schemes and Tannaka categories.

Theorem (1.2) : A group-scheme G is finite if and only if there exists a finite collection S of G-representations which generates G-mod (in the sense of the above definition).

Theorem (1.3) : Any homomorphism of Tannaka categories from (G-mod, $\hat{\otimes}$ T_k, L_0) to (H-mod, $\hat{\otimes}$, T_k, L_0) is induced by a unique homomorphism (of affine algebraic group schemes) from H to G.

2.2. *Principal bundles*

Let X be a nonempty k-prescheme, $S(X)$ the category of quasi-coherent sheaves on X, $\otimes : S(X) \times S(X) \to S(X)$ the tensor product functor on sheaves.

Let G be an affine group scheme defined over k.

Recall that $j : P \to X$ is said to be a principal G-bundle on X if

(a) j is a surjective flat affine morphism.

(b) $\Phi : P\ G \to P$ defines an action of G on P such that $j \cdot \bar{\phi} = j \cdot p_1$.

(c) $\Psi : P \times G \to P \times_X P$ by $\Psi = (p, \Phi)$ is an isomorphism.

In this case, $\mathscr{F} \to j^* (\mathscr{F})$ gives an isomorphism of $S(X)$ with the category of G-sheaves on P, by the method of flat descent (see Grothendieck [2]). Every left representation V of G gives rise to a G-sheaf on P in a natural way, and by taking G-invariants, one gets a sheaf on X, denoted by $F(P)\,V$. This gives rise to a functor $F(P) : G\text{-mod} \to S(X)$, and putting $F = F(P)$, we note that the following are true :

F_1 : F is a k-additive exact functor ; F_2 : $F \circ \hat{\otimes} = \otimes \circ (F \times F)$; F_3 : The obvious statements parallel to $\mathcal{C}5$, $\mathcal{C}6$, $\mathcal{C}7$; in particular, $FL_0 = \mathbf{O}_X$, where L_0 is the trivial representation, and finally ; F_4 : If rank $V = n$, then FV is locally free of rank n ; in particular, F is faithful.

From now on, F will denote a functor where $F1$ to $F4$ held.

Let G-mod be the category of all (possibly infinite-dimensional) left representations of G.

Lemma (2.1) : There is a unique functor $\bar{F} : G\text{-mod} \to S(X)$, such that :

(i) The statements $F1$, $F2$, $F3$ hold good for $\bar{F}$, (ii) $\bar{F} \mid G\text{-mod} = \bar{F}$, (iii) $\bar{F}V$ is flat for all V, and faithfully flat if $V \neq 0$, and (iv) $\bar{F}$ preserves direct limits.

Proof : Define $\bar{F}V$ to be the direct limit of FW, where W runs through the collection of finite-dimensional G-invariant sub-spaces of V, and the lemma is then easily checked. We will put $\bar{F} = F$ from now on.

Lemma (2.2) : F induces a functor from affine G-schemes to affine X-preschemes.
Proof : Let $Y = \text{spec } A$ be a scheme on which G operates, and let $m : A \otimes A \to A$ be the multiplication map on A. Since A is a commutative, associative k-algebra with identity, by $F2$ and $F3$, we deduce that FA is a commutative, associative sheaf of $\mathbf{O}_X$-algebras with identity. This is enough to conclude that

there is an affine morphism $j : Z \to X$ such that $j_*(\mathbf{O}_Z)$ is isomorphic to FA as a sheaf of $\mathbf{O}_X$-algebras. We shall denote Z by FY from now on.

Definition : Let G operate on itself by the left. Put $P(F) = FG$, and let $j : P(F) \to X$ be the canonical morphism. Since no confusion is likely to arise, we shall denote $P(F)$ simply by P.

Lemma (2.3) : P is a principal G-bundle on X.
Proof : By definition, j is an affine morphism. That j is flat and surjective follows from the fact that $j_*(\mathbf{O}_P)$ is faithfully flat. ((iii) of Lemma 2.1). Properties (*b*) and (*c*) will be checked later.
Lemma (2.4) : If Y and Z are schemes on which G operates, $F(Y \times Z) = FY \times_X FZ$. Furthermore, if G acts trivially on Y, then $FY = X \times Y$.
Proof : Obvious.

Proof of Lemma (2.3) : We denote by G' the same scheme as G, equipped with the trivial action of G. Let $\varphi : G \times G' \to G$ be the multiplication map of G, and $\psi : G \times G' \to G \times G$ be given by $\psi(x, y) = (x, \varphi(x, y))$. Note that φ and ψ are both G-morphisms ; consequently there are X-morphisms

$$\Phi = F\varphi : P \times G \to P, \text{ and}$$

$$\Psi = F\psi : P \times G \to P \times_X P.$$

Since φ defines an action of G' on G, Φ defines an action of G on P, and $j \circ p_1 = j \circ \bar{\Phi}$ simply because Φ is an X-morphism.

Also, ψ is an isomorphism, from which it follows that Ψ is an isomorphism too, thus concluding the proof of the lemma.

Now that we have constructed a principal bundle P, given a functor F, the next step is to show that F is the functor naturally associated with P, that is:

Proposition (2.5) : $F = F(P)$.
We introduce some notation first. Let Z be a scheme on which G operates on the right, and let V be any left representation of G. We denote by V_Z the sheaf $V \otimes \mathbf{O}_Z$ equipped with the following action of $G : g(v \otimes f) = gv \otimes f \circ \rho(g)$, where $v \in V$, $g \in G$, and $f \in \Gamma(U, \mathbf{O}_Z)$, for some open U in Z. This is the natural construction of a G-sheaf on Z, given a representation V, mentioned in the beginning of the section.

To show that two sheaves are isomorphic on X, it suffices to prove that the inverse images are isomorphic as G-sheaves on P, and hence the above proposition is reduced to the following :

Lemma (2.6) : There is a functorial isomorphism (of G-sheaves) of $j^*(FV)$ with V_P.
We require the aid of

Lemma (2.7) : Let Y be an affine scheme on which G operates on the left, H operates on the right. Assuming the actions of G and H on Y commute with each other, let $Z = FY$. Then Z is a H-scheme, and $j : Z \to X$ is a H-morphism, where X has the trivial action of H. Furthermore, F induces a functor F

from the category of sheaves on Y with commuting G and H action to the category of H sheaves on Z.

The proof of Lemma 2·7 is trivial, and we omit it. To apply the Lemma, put $G = H = Y$, with the actions of G and H on Y being given by left and right translations respectively.

Let V be a representation of G and V' its underlying vector space equipped with the trivial action of G. Therefore, there are G-sheaves, V_G and V'_G (corresponding to the right action of G) on G. We shall define left actions of G on V_G and V'_G as follows :

(a) $g(v \otimes f) = v \otimes f \circ L_g^{-1}$, for $v \in V, g \in G$, and $f \circ \Gamma(U, \mathbf{O}_G)$,

(b) $g(v \otimes f) = gv \otimes f \circ L_g^{-1}$, for $v \in V'$, $g \in G$, and $f \circ \Gamma(U, \mathbf{O}_G)$.

With F as in Lemma 2·7, it is trivial to check that $F(V_G) = V_P$ and $F(V'_G) = j^*(FV)$. To prove Lemma 2·6, it therefore suffices to prove

Lemma (2·8) : There is a functorial isomorphism of V with V'_G as sheaves on G, with G acting both on the left and the right.

Proof : Let W be any vector space. We denote the scheme spec $(S(W^*))$ again by W. Then the sheaf $W \otimes \mathbf{O}_G$ can be identified canonically with the sheaf of morphisms from G to the scheme W.

Using this identification, define $\lambda : V'_G \to V_G$ by $\lambda(f)(g) = g^{-1}f(g)$, where $g \in G, f : G \to V'$. The map furnishes the required isomorphism, thus concluding the proof of Prop. 2·5.

Proposition (2·9) : There is a bijective correspondence between principal G-bundles on X and functors $F : G\text{-mod} \to \mathcal{S}(X)$ such that $F1$ to $F4$ hold. Furthermore,

(a) Let $f : Y \to X$ be a morphism, and assume that $F = G\text{-mod} \to S(X)$ satisfies $F1$ to $F4$. Then $F1$ to $F4$ hold good for $f^* \circ F$ also, and $P(f^* \circ F) = f^*(P(F))$; (b) Let $X = \text{spec}\, k$, and $F : G\text{-mod} \to k\text{-mod}$ the forgetful functor. Then $P(F) = G$; (c) Let $\rho : H \to G$ be a morphism of affine group schemes. Let P be a principal H-bundle on X, and P' the quotient of $P \times G$ by H. Let R : G-mod $\to$ H-mod be the restriction functor. Then $F(P) \circ R = F(P')$.

Proof : (b) is trivial, and (a) and (c) are proved by chasing the construction of $P(F)$.

Remark : The condition $F4$, which is crucial in proving that $j : P \to X$ is flat and surjective, is actually a consequence of $F1$, $F2$ and $F3$. However, we do not need this fact.

2·3. *Essentially finite vector bundles*

Let X be a complete connected reduced k-scheme, where k is a perfect field. Let vect(X) denote the set of isomorphism classes, $[V]$, of vector bundles V, on X. Then vect(X) has the operations :

(a) $[V] + [V'] = [V \otimes V']$, and

(b) $[V] \cdot [V'] = [V \times V']$.

In particular, for any vector bundle V on X, given a polynomial f with non-negative integer coefficients, $f(V)$ is naturally defined.

Let $K(X)$ be the Grothendieck group associated to the additive monoid vect (X); this is not the usual Grothendieck ring of vector bundles on X, since $0 \to V' \to V \to V'' \to 0$ exact does not imply that $[V'] + [V''] = [V]$.

The Krull–Schmidt–Remak theorem holds, since $H^\circ(X, \text{end } V)$ is finite-dimensional. In particular, $[W]$, where W runs through all indecomposable vector bundles on X, form a free basis for $K(X)$.

Definition : For a vector bundle V, $S(V)$ is the collection of all the indecomposable components of $V^{\otimes n}$, for all non-negative integers n.

Lemma (3.1) : Let V be a vector bundle on X. The following are equivalent :

(a) $[V]$ is integral over $\mathbf{Z}$ in $K(X)$; (b) $[V] \otimes 1$ is integral over Q in $K(X) \otimes Q$; (c) There are polynomials f and g with non-negative integer coefficients such that $f(V)$ is isomorphic to $g(V)$ and $f \neq g$.

(d) $S(V)$ is finite.

Proof :

(a) $\Leftrightarrow$ (b) holds merely because $K(X)$ is additively a free abelian group.

(c) $\Rightarrow$ (b) is obvious.

(b) $\Rightarrow$ (c) : Let $h \in \mathbf{Z}[t]$ such that $h([V]) = 0$, and $h \neq 0$. Choose $f, g \in \mathbf{Z}[t]$ such that f and g have non-negative coefficients, and $h = f - g$. Then $[f(V)] = [g(V)]$ in $K(X)$, but vect (X), as a monoid, has the cancellation property, so it follows that $f(V)$ is actually isomorphic to $g(V)$.

(d) $\Rightarrow$ (a) : The abelian subgroup of $K(X)$ with basis as $S(V)$ is certainly stable under multiplication by $[V]$.

(a) $\Rightarrow$ (d) : Simply note that if m is the degree of a monic polynomial h such that $h([V]) = 0$, then any member of $S(V)$ is actually an indecomposable component of $V^{\otimes r}$ for some r lying between 0 and $m - 1$.

Definition : A vector bundle V on X is said to be finite if it satisfies any of the equivalent hypothesis of Lemma 3·1.

Lemma (3.2) :

(i) V_1, V_2 finite $\Rightarrow$ $V_1 \oplus V_2$, $V_1 \otimes V_2$, V_1^* finite.

(ii) $V_1 \oplus V_2$ finite $\Rightarrow$ V_1 finite.

(iii) A line bundle L is finite $\Leftrightarrow$ $L^{\otimes m}$ is isomorphic to $\mathbf{O}_X$ for some positive integer m.

Proof :

(1) is obvious.

(2) follows from the fact that $S(V_1)$ is contained in $S(V_1 \oplus V_2)$.

(3) follows from the fact that $S(L) = \{L^{\otimes m} : m \geqslant 0\}$.

Lemma (3.3) : Let X be a smooth projective curve. For a vector bundle V, let $C(V) = \sup\{\mu(W) = \deg W/rkW,\ 0 \neq W \subseteq V\}$.

Then, (a) $C(V)$ is finite, and

(b) if $0 \to V' \to V \to V'' \to 0$ is an exact sequence of vector bundles on X, $C(V) \leqslant \max(C(V'), C(V''))$.

Proof : That $D(V) = \sup\{\deg L : L \subset V,\ L$ a line bundle$\}$ is finite is well-known. Since $C(V) \leqslant \max\{D(\Lambda^r(V))/r : 1 \leqslant r \leqslant rk\, V\}$, (a) follows.

Given an injection $j : W \to V$ and an exact sequence $0 \to V' \to V \to V'' \to 0$, there is a canonical factoring :

$$\begin{tikzcd} 0 \arrow[r] & W' \arrow[r] \arrow[d, "j'"'] & W \arrow[r] \arrow[d, "j"'] & W'' \arrow[r] \arrow[d, "j''"'] & 0 \\ 0 \arrow[r] & V' \arrow[r] & V \arrow[r] & V'' \arrow[r] & 0 \end{tikzcd}$$

such that the horizontal rows are exact, and j', j'' are generic injections. Let U', U'' be the sub-bundles of V', V'' respectively, such that $j'(W') \subseteq U'$ and $j''(W'') \subseteq U''$, and $rkW' = rkU'$ and $rk\,W'' = rk\,U''$. Then $\deg W' \leqslant \deg U'$ and $\deg W'' \leqslant \deg U''$.

Now,

$$\begin{aligned} \mu(W) &= \deg W' + \deg W''/rk\ W' + rk\ W'' \\ &\leqslant \deg U' + \deg U''/rk\ U' + rk\ U'' \\ &\leqslant \max(\deg U'/rk\ U', \deg U''/rk\ U'') \\ &\leqslant \max(C\,V'), C(V''), \end{aligned}$$

which proves (b).

Proposition (3.4) : Any finite vector bundle V on a smooth projective curve X is semistable of degree zero.

Proof : By Lemma 3.3, $C(V^{\otimes m}) \leqslant \sup\{C(W) : W \in S(V)\} = T(V)$, which is finite, since $S(V)$ is a finite collection. Consequently, for any sub-bundle W of V, $W \neq 0$, since $W^{\otimes m}$ is a sub-bundle of $V^{\otimes m}$, $\mu(W^{\otimes m}) \leqslant T(V)$, for all non-negative integers m. But a simple calculation shows that $\mu(W^{\otimes m}) = m\mu(W)$ which obviously implies that $\mu(W) \leqslant 0$.

In particular, since both V and V^* are finite, $\mu(V) \leqslant 0$ and $\mu(V^*) = -\mu(V) \leqslant 0$. Therefore we have shown that

(a) $\mu(V) = 0$, and

(b) for all sub-bundles W of V, $W \neq 0$, $\mu(W) \leqslant 0$.

For the rest of this section, X will be a complete, connected, reduced scheme and the phrase "a curve Y in X" is to be interpreted as a morphism $f : Y \to X$, where Y is a smooth, connected, projective curve, and f is a birational morphism onto its image.

Definition : A vector bundle on X is semistable if and only if it is semistable of degree zero restricted to each curve in X.

Since the restriction of a finite vector bundle is also finite, we have the following obvious corollary :

Corollary (3.5) : A finite vector bundle on X is semistable.

Lemma : (3.6) :

(a) If V is a semistable vector bundle on X and W is either a sub-bundle or a quotient bundle of V, such that W/Y has degree zero for each curve Y in X, then W is semistable.

(b) The full subcategory of $\mathcal{S}(X)$ with objects as semistable vector bundles on X is an abelian category.

Proof :

(a) Under the given hypothesis, it follows that $W \mid Y$ is semistable of degree zero, and therefore W is semistable.

(b) Let V and W be semistable vector bundles on X, and let $f : V \to W$ be a morphism. For a geometric point x : spec $\bar{k} \to X$, let $r(x)$ be the rank of the morphism $x^*(f) : x^*(V) \to x^*(W)$. Then, by elementary degree considerations $r(x)$ is a constant restricted to each curve, and since X is connected, $r(x)$ is constant globally. Now, since X is reduced, it follows that ker f and coker are locally free, and moreover, $(\ker f)|Y = \ker(f \mid Y)$ and $(\operatorname{coker} f)|Y = \operatorname{coker}(f/Y)$, and both these bundles are semistable of degree zero on Y; the lemma follows.

Definition : We shall denote by $SS(X)$ the full subcategory of $S(X)$ with semistable vecor bundles as objects. Let F be the collection of finite vector bundles, regarded as a subset of obj $SS(X)$, and let $EF(X)$ be the full subcategory of $SS(X)$ with obj $EF(X) = \bar{F}$, where the meaning of $\bar{F}$ is to be taken in the sense of § 1. The objects of $EF(X)$ will be essentially called finite vector bundles.

Proposition (3·7) :

(a) If V is an essentially finite vector bundle on X, and W is either a sub-bundle or a quotient bundle of V such that $W \mid Y$ has degree zero for each curve Y in X, then W is essentially finite. (b) $EF(X)$ is an abelian category. (c) If V_1 and V_2 are essentially finite, so are $V_1 \otimes V_2$ and V^*.

Proof : (a) and (b) are obvious consequences of Lemma 3·6. To prove (c), choose W_i and P_i such that

(i) W_i is finite, (ii) P_i is a sub-bundle of W_i and P_i is semistable, and (iii) V_i a quotient of P_i, for $i = 1, 2$.

Then

(i) $W_1 \otimes W_2$ is finite by Lemma 3.2, (ii) $P_1 \otimes P_2$ is a sub-bundle of $W_1 \otimes W_2$, and (iii) $V_1 \otimes V_2$ is a quotient of $P_1 \otimes P_2$.

Both $P_1 \otimes P_2$ and $V_1 \otimes V_2$ are of degree zero restricted to each curve in X; consequently, by (a), both $P_1 \otimes P_2$ and $V_1 \otimes V_2$ are essentially finite.

In a similar fashion, one proves that the dual of an essentially finite vector bundle is essentially finite.

Proposition (3·8) : Let G be a finite group scheme, and $j : X' \to X$ a principal G-bundle. Then, for the functor $F(X') : G\,\mathrm{mod} \to S(X)$, $F(X')\,V$ is always an essentially finite vector bundle.

Proof : We shall show that $F(X')\cdot V$ is of degree zero restricted to each curve. For this, we assume that X itself is a smooth projective curve. Let R be the coordinate ring of G, and n the vector space dimension of R. Then, n deg $(F(X')\,V) = \deg(j^*(F(X')\,V)$; but $j^*(F(X')\,V)$ is, by definition, a trivial vector bundle on X', and therefore deg $(F(X')\,V)$ is equal to zero. Note that "degree" makes sense even if X' is ot reduced, by looking at Hilbert polynomials.

Now, any representation V of G can be embedded (injectively) in $R \oplus R \cdots \oplus R$, and therefore $F(X') V$ is contained in a direct sum of several copies of $F(X') R$. To prove that $F(X') V$ is essentially finite, it would surffice to show that $F(X') R$ is finite, by (a) of Prop. 3·7. But $R \otimes R$ is isomorphic to $R \oplus R \oplus \ldots \oplus R$ n times, from which, if $W = F(X') R$, $[W]^2 = n[W]$, concluding the proof of the proposition.

For the rest of this section, we shall fix a k-rational point x of X and denote by $x^* : S(X) \to |k|$ the functor which associates to a sheaf on X its fibre at the point x. Note that x^* is faithful and exact when restricted to the category of semistable bundles. It is now obvious that $(EF(X), \otimes, x^*, O_X)$ is a Tannaka category. By Theorem 1·1, this determines an affine groups scheme G such that $|G|$ can be identified with $EF(X)$ in such a way that x^* becomes the forgetful functor. We shall call the group-scheme G above the fundamental group-scheme of X at x, and denote it by $\pi(X, x)$.

For a subset S of obj $EF(X)$, let $S^* = \{V^* : V \in S\}$. Let $S_1 = S \cup S^*$ and $S_2 = \{V_1 \otimes V_2 \otimes \cdots \otimes V_m : V_i \in S_1\}$. Let obj $EF(X, S) = \bar{S}_2$. As before, this determines an affine group scheme which we call $\pi(X, S, x)$, such that

$$G_S : EF(X, S) \to |\pi(X, S, x)|^*$$

is an equivalence of categories. Let F_S be the inverse of G_S; then F_S can be regarded as a functor from $|\pi(X, S, x)|$ to $S(X)$ such that the composite x^*. F_S is the forgetful functor. In particular, by Prop. 2.9, there is a principal $\pi(X, S, x)$-bundle $\tilde{X}_S$ such that $F_S = F(\tilde{X}_S)$. By Prop. 2·9 (a), the functors $x^* \cdot F_S$ and $F(\tilde{X}_S \mid x)$ coincide, and by Prop. 2.9 (b), there is a natural isomorphism of $\tilde{X}_S \mid x$ with G (as G-spaces), which is equivalent to specifying a k-rational base point $\bar{x}_s$ of $X_S \mid x$.

Now, if S is a subset of Q, there is a natural homomorphism of Tannaka categories from $EF(X, S)$ to $EF(X, Q)$, which by Theorem 1.3, determines a natural homomorphism ρ^Q_S from $\pi(X, Q, x)$ to $\pi(X, S, x)$, and by Prop. 2.9 (c), it follows that X_S is induced from X_Q by the homomorphism ρ^Q_S.

Lemma (3·9) : Let S be a finite collection of finite vector bundles. Then $\pi(X, S, x)$ is a finite group scheme.

Proof : Let W be the direct sum of all the members of S and their duals. Then W is a finite vector bundle, and by Lemma 3.1, $S(W)$ is finite. Note that $S(W)$ generates the abelian category $EF(X, S)$ in the sense of § 1, and therefore, by Theorem 1.2, $\pi(X, S, x)$ is a finite group-scheme.

Proposition (3 .10) : Let S be any finite collection of essentially finite vector bundles. Then, there is a principal G-bundle X' on X, with G a finite groud scheme, such that the image of $F(X')$: G and $\to S(X)$ contains the given collection S.

Proof : For each $W \in S$, choose V such that W is a quotient of a semistable sub-bundle of V and V a finite vector bundle. Let Q be the collection of V as constructed, and note that S is a subset of obj $EF(X, Q)$.

Put $G = \pi(X, Q, x)$, and $X' = \tilde{X}_Q$. By Lemma 3.9, G is a finite group-scheme. Let G_Q, as above, be the equivalence of categories, from $EF(X, Q)$ to $|\pi(X, Q, x)|$, and then, we know that $F(X')$. $G_Q(V) = V$, for all objects V of $EF(X, Q)$, thus proving the proposition.

* $|\pi(X, S, x)|$ denote the cotegory of $\pi(X, S, x)$-modules.

For $S = \text{obj}\, EF(X)$, we shall denote $\tilde{X}_S$, G_S, $\tilde{F}_S$, $\tilde{x}_S$ by $\tilde{X}$, G, $\tilde{F}$, $\tilde{x}$ respectively.

Definition : The principal $\pi(X, x)$-bundle $\tilde{X}$ is the universal covering scheme of X.

The universal property possessed by $\pi(X, x)$ and $\tilde{X}$ is given by the following :

Proposition (3.11) : Let (X', G, u) be a triple, such that X' is a principal G-bundle on X, u a k-rational point in the fibre of X' over X, and G is a finite group scheme.

Then there is a unique homomorphism $\rho : \pi(X, x) \to G$, such that

(a) X' is induced from $\tilde{X}$ by the homomorphism ρ, and

(b) the image of $\tilde{x}$ in X' is u.

Consequently, there is a bijective correspondence of the above triples with homomorphisms $\rho : \pi(X, x) \to G$.

Proof : By Prop. 3.8, $F(X')$ is a functor from mod-G to $EF(X)$. Now $EF(X)$ is identified with $|\pi(X, x)|$ in such a way that the forgetful functor T_k on $|\pi(X, x)|$ is equivalent to the functor x^* from $EF(X)$ to $|k|$. Thus, the composite $T_k \cdot F(X')$ is simply $x^* \cdot F(X') = F(X' | x)$, by Prop. 2·9 (a). Now, the k-rational point u of $X' | x$ gives a unique isomorphism $\varphi : G \to X' | x$ of principal homogeneous spaces such that $\varphi(1) = u$. By Prop. 2·9 (b), φ determines a natural equivalence of the functor $F(X' | x)$ with the forgetful functor from mod-$G \to$ mod-k. This information yields a morphism (of Tannaka categories) from mod-G to $\pi(X, x) |$, which, by Theorem 1·3, is induced by a homomorphism $\rho : \pi(X, x) \to G$. We now appeal to Prop. 2·, (c) to settle the fact that X' is indeed induced from $\tilde{X}$ by ρ, and that the image of $\tilde{x}$ in X' is u. The uniqueness of ρ is easily checked.

3. Conclusion

(1) With S as in Lemma 3·9, assume that $|\pi(X, S, x)|$ is a semi-simple category. Then, for any representation W of $\pi(S, X, x)$, there exist polynomials f and g, with $f \neq g$, the coefficients of f and g being non-negative integers, such that $f(W)$ and $g(W)$ are isomorphic. This follows from the fact that there are only finitely indecomposable representations of $\pi(X, S, x)$ up to isomorphism. Putting $V = F(\tilde{X}_S)\, W$, it follows that V is a finite vector bundle.

In characteristic zero, any finite group scheme is reduced, and its representations certainly form a semi-simple category. By Prop. 3·10, it follows therefore that in characteristic zero, any essentially finite vector bundle is finite.

(2) The structure of the fundamental group-scheme :

(a) For $S \subseteq Q \subseteq \text{obj}\, EF(X)$,

$$\rho_S^Q : \pi(X, Q, x) \to \pi(X, S, x) \text{ is surjective.}$$

(b) $\pi(X, x)$ is the inverse limit of $\pi(X, S, x)$, where S runs through all finite collections of finite bundles on X ; consequently $\pi(X, x)$ is the inverse limit of finite group schemes.

Both (a) and (b) follow from standard facts about Tannaka categories (Saavedra Rivano 1972).

CHAPTER II

The fundamental group-scheme

We shall always assume that the base-field k has characteristic $p > 0$.

Propositions 1 and 2 examine the existence of the fundamental group-scheme of a k-scheme X with a k-rational base-point χ_0. The results of Chapter I are re-interpreted in Proposition 3. Proposition 4 studies the dependence of $\pi(X, \chi_0)$ on the base-point χ_0. Proposition 5 shows that $\pi(X, \chi_0)$ is well behaved under base-change by separable algebraic extensions of the ground field. This is the content of § 1.

In § 2 we show that the fundamental group-scheme is a birational invariant for smooth complete varieties. This involves a "purity of branch-locus" theorem (see Proposition 7). Finally we show that $\pi(X, \chi_0)$ is trivial for normal complete rational varieties.

§ 1. X is a k-scheme and $\chi_0 : \mathrm{spec}\, k \to X$ is a morphism.

Consider the following category $\mathcal{C}$: the objects are triples (Q, G, V) where Q is a principal G-bundle on X, G is a finite group-scheme and V is a k-rational point of Q sitting above χ_0.

A morphism $(f, g) : (Q, G, V) \to (Q', G', V')$ is a homomorphism $g : G \to G'$ and a morphism $f : Q \to Q'$ that intertwines the G and G' actions on Q and Q' and in addition, $f(V) = V'$.

By $\mathcal{C}^1$ we shall denote the category of all triples (Q, G, V) as above, except that G is now an inverse limit of finite group-schemes.

Definition 1 : X has a fundamental group-scheme $\pi(X, \chi_0)$ if there is a triple $(P, \pi(X, \chi_0), *)$ in the category $\mathcal{C}'$ such that for each object (Q, G, V) of $\mathcal{C}'$, there is a unique morphism from $(P, \pi(X, \chi_0), *)$ to (Q, G, V).

Remark : Clearly it suffices to check the above for all (Q, G, V) in $\mathcal{C}$ to ensure that it holds for all (Q, G, V) in $\mathcal{C}^1$.

Definition 2 : X has $\mathcal{P}$ if whenever $(f_i, \rho_i) : (Q_i, G_i, V_i) \to (Q, G, V)$ for $i = 1, 2$ are morphisms in $\mathcal{C}$, then

$$(Q_1 \times_Q Q_2\ G_1 \times_G G_2,\ v_1 \times v_2) = (Z, H, w) \text{ is an object of } \mathcal{C}.$$

We have :

Proposition 1 : X has a fundamental group-scheme if and only if X has $\mathcal{P}$.

Proposition 2 : If X is reduced, and connected, then X has a fundamental group-scheme.

First we need

Lemma 1 : With $Q_i, G_i, V_i, \rho_i, Z, H, w$ as in Definition 2, Z is a principal H-bundle on a closed sub-scheme R of X containing x_0.

Proof of Lemma 1

$Z = Q_1 \times_Q Q_2$ is a closed sub-scheme of $T = Q_1 \times_X Q_2$ which is a principal $(G_1 \times G_2)$-bundle on X. Let q_i be the composite $T \to Q_i \xrightarrow{f_i} Q$ for $i = 1, 2$. Because Q is a principal G-bundle on X, there is a unique $z : T \to G$ such that $q_1 = q_2\, z$. If $e \to G$ represents the identity of G, clearly $Z = z^{-1}(e)$.

Also there is a commutative diagram :

$$\begin{tikzcd} T \times G_1 \times G_2 \arrow[r] \arrow[d, "z \times 1_{G_1} \times 1_{G_2}"'] & T \times_X T \arrow[d, "z \times z"] \\ G \times G_1 \times G_2 \arrow[r, "h"] & G \times G \end{tikzcd}$$

where the first horizontal arrow induces $(p, g) \to (p, pg)$ for $p \in \mathrm{Mor}(S, T)$, $g \in \mathrm{Mor}(S, G_1 \times G_2)$, and h induces $(g, g_1, g_2) \to (g, \rho_2(g_2)^{-1} g \rho_1(g))$ for all $g_1 \in \mathrm{Mor}(S, G_1)$, $g_2 \in \mathrm{Mor}(S, G_2)$ and $g \in \mathrm{Mor}(S, G)$.

Now $h^{-1}(e \times e) = e \times H$, and taking inverse images in the vertical arrows, we see that the first horizontal arrow restricts to an isomorphism $Z \times H \to Z x_x Z$.

In particular Z is stable under the H-action on T; this action, being free, makes Z a principal H-bundle on $R = Z/H$ which is a closed sub-scheme of T/H. only remains to show that $R \to X$ is a closed immersion.

Consider the commutative diagram :

$$\begin{tikzcd} Z \times H \arrow[r, "\cong"] \arrow[d, "j \circ p_1"'] & Z \times_X Z \arrow[d, "j \times j"] \\ R \arrow[r, "\Delta"] & R \times_X R \end{tikzcd}$$

where $j : Z \to R$ is the given morphism and $p_1 : Z \times H \to Z$ is the projection. The morphism $j \times j$ makes $zx_X Z$ a principal bundle on $Rx_X R$ with the obvious action of $H \times H$ on the right. The right action of $H \times H$ on $Z \times H$ given by $(p, h) \cdot (h_1, h_2) = (ph_1, h_1^{-1} h h_2)$ for all $p \in \mathrm{Mor}(S, Z)$, h, h_1 and $h_2 \in \mathrm{Mor}(S, H)$ clearly makes $Z \times H$ a principal bundle on R. Moreover $Z \times H \to Z x_X Z$ preserves the $H \times H$ action. It follows that $\triangle$ is an isomorphism.

It is a trivial matter to check that a finite morphism $A \to B$ is a closed immersion if and only if the diagonal $A \xrightarrow{\Delta} A x_B A$ in an isomorphism. So this proves that $R \to X$ is a closed immersion, finishing the proof of the lemma.

With z as above $W = Z^{-1}(G_{\mathrm{loc}})$ is an open and closed sub-scheme of T. But $\pi : T \to X$ is flat implying that $\pi(W) = X$ if X is connected. We have seen that $z^{-1}(e) = Z$, from which it follows that $\pi(Z)_{\mathrm{red}} = X_{\mathrm{red}}$, and therefore $\pi(Z) = X$ if X is assumed to be reduced. Consequently $R = X$ and this shows that Z is a principal H-bundle on X. Thus any connected and reduced X has $\mathscr{P}$.

Next we show that if X has a fundamental group-scheme, then X has $\mathscr{P}$. By definition there exist $(r_i, s_i) : (P, \pi(X, X_0), *) \to (Q_i, G_i, V_i)$ for $i = 1, 2$, and by uniqueness $(f_1 r_1, \rho_1 \circ s_1) = (f_2 r_2, \rho_2 \circ s_2)$. Therefore $r_1 \times r_2 : P \to T$ has $(r_1 \times r_2) P \subseteq Z$. This shows that $R = \pi(Z) = X$ and therefore Z is a principal H-bundle on X, as was to be shown.

Finally we show that if X has $\mathscr{P}$, then X has a fundamental group-scheme. First some generalities :

A small category $\mathscr{D}$ is an inverse system if given $f_i : A_i \to B$ for $i = 1, 2$, there is an object C of $\mathscr{D}$ and morphisms $g_i : C \to A_i$ for $i = 1, 2$ that $f_1 \circ g_1 = f_2 \circ g_2$.

Given such a category and a functor $F : \mathscr{D} \to \mathscr{C}'$, there is a canonically associated object of $\mathscr{C}'$.

For any object A of $\mathfrak{D}$, if $FA = (Q, G, V)$, put $QFA = Q$, $GFA = G$ and $VFA = V$. The triple $(\tilde{Q}, \tilde{G}, \tilde{V})$ is an object of $\mathcal{C}'$:

$$\tilde{Q} = \varprojlim_{A \in \mathfrak{D}} QFA, \quad \tilde{G} = \varprojlim_{A \in \mathfrak{D}} GFA, \quad \tilde{V} = \varprojlim_{A \in \mathfrak{D}} VFA.$$

Let us check that these constructions make sense :

If $R(G)$ denotes the coordinate ring of an affine group-scheme G, then $R = \varinjlim_{A \in \mathfrak{D}} R(GFA)$ is a Hopf algebra which is the union of its finite dimensional Hopf sub-algebras. Therefore $\tilde{G} = \operatorname{spec} R$ is an inverse limit of finite group-schemes.

Similarly, if $R(Q) = j_*(\mathbf{O}_Q)$ for a morphism $j : Q \to X$, put $R' = \varinjlim_{A \in \mathfrak{D}} R(QFA)$ is a locally free sheaf of $\mathbf{O}_X$-algebras on X ; thus there is a flat affine morphism $j : \tilde{Q} \to X$ such that $R(\tilde{Q}) = R'$.

The isomorphisms $QFA \otimes_k GFA \to QFA \otimes_X QFA$ give isomorphisms $R(QFA) \otimes_X R(QFA) \to R(QFA) \otimes_k R(GFA)$, and in the direct limit ar isomorphism $R' \otimes_X R' \to R' \otimes_k R$. Thus $Q \times G \to Q \times_X Q$ is an isomorphism. $\tilde{V}$ is constructed similarly.

If we assume that X has $\mathcal{P}$, then $\mathcal{C}$ itself is an inverse system and the $(\tilde{Q}, \tilde{G}, \tilde{V})$ associated to the identity functor of $\mathcal{C}$ is easily seen to satisfy all the required properties of $(P, \pi(X, x_0), *)$.

This completes the proofs of Propositions 1 and 2.

In Chapter I, we constructed $\pi(X, x_0)$ directly using the Tannaka category of all essentially finite vector bundles. The essential content of Proposition 3·11 is contained in Proposition 3 given below.

Definition 3 : A triple (Q, G, V) in $\mathcal{C}$ is reduced if for any morphism $(Q', G', V') \to (Q, G, V)$ in $\mathcal{C}$, $G' \to G$ is a surjection (i.e., it is a surjection in the flat topology : more directly $R(G) \to R(G')$ is an injection).

If X has a fundamental group-scheme, a triple (Q, G, V) is reduced if and only if the homomorphism $\pi(X, \chi_0) \to G$ is surjective. This is trivial.

Proposition 3 : Let X be a complete, connected and reduced k-scheme with χ_0 as usual. Let (Q, G, V) be a triple in $\mathcal{C}$. Then the following are equivalent :

(a) (Q, G, V) is reduced.

(b) The functor $F(Q)$: mod-$G \to S(X)$ is fully faithful (see § 2, Chapter I for the definition of $F(Q)$).

(c) $\Gamma(Q, \mathbf{O}_Q) = k$.

This is the only case where we have a criterion for determining whether (Q, G, V) is reduced or not. In characteristic zero, (Q, G, V) is reduced if and only if Q is connected, as is well-known.

Proof: $B \Rightarrow C$. Let $j : Q \to \mathscr{X}$ be the given morphism. Then $\Gamma(Q, O_Q) = \Gamma(\mathscr{X}, j_* O_Q) = \Gamma(\mathscr{X}, F(Q)R(G))$ = the fixed subspace of $R(G)$ under the G-action (because $F(Q)$ is fully faithful) $= k$.

$C \Rightarrow A$. There is a morphism $(f, \rho) = (P, \pi(\mathscr{X}, \chi_0), *) \to (Q, G, V)$. If ρ is not surjective, its image is a proper closed subgroup-scheme H of G and the fixed subspace of $R(G)$ under the $\pi(\mathscr{X}, \chi_0)$-action is clearly the coordinate ring of G/H which contains k properly. Therefore $\Gamma(Q, O_Q) = \Gamma(\mathscr{X}, F(P)R(G))$ contains k properly.

$A \Rightarrow B$. We are given that $\rho : \pi(\mathscr{X}, \chi_0) \to G$ is a surjective. Therefore G-mod $\to \pi(\mathscr{X}, \chi_0)$-mod is fully faithful. But from the construction of $\pi(\mathscr{X}, \chi_0)$ in chapter I, $\pi(\mathscr{X}, \chi_0)$-mod $\to S(\mathscr{X})$ is fully faithful. Thus mod-$G \to S(\mathscr{X})$ is fully faithful.

Next we deal with the relation between $\pi(\mathscr{X}, \chi_0)$ and $\pi(\mathscr{X}, \chi_1)$ where χ_1 : spec $k \to \mathscr{X}$ is another k-rational point of $\mathscr{X}$, assuming that $\pi(\mathscr{X}, x_0)$ does indeed exist

Let R be a principal homogeneous G-space with G acting on the left. Then there is a group-scheme G' acting on R on the right such that

(a) the actions of G and G' on R commute, and

(b) R is a principal homogeneous G'-space.

This determines G' uniquely. In the literature, G' is often called an inner twist of G, especially when G is an affine algebraic group. If G is commutative, then $G = G'$.

Let $\mathscr{C}_1$ be the same category as $\mathscr{C}$ except that the base-point V of Q sits above χ_1.

Take an object (Q, G, V) of $\mathscr{C}$. Let R and G' be as above. Then $Q \times R$ has an action of $G \times G'$: G' acts trivially on Q and on R in the given manner, and $Q \times R$ gets the diagonal action of G. The quotient $Q' = Q \times R/G$ is thus a principal G'-bundle on $\mathscr{X}$.

In particular, $j^{-1}(\chi_1)$ in $j : Q \to \mathscr{X}$ is a principal homogeneous G-space, so we may put $R = j^{-1}(\chi_1)$. In this case the fibre over χ_1 in Q' is $R \times R/G$ which contains $\Delta R/G$. This $\Delta R/G$ gives a base point V' of Q' sitting above χ_1. Thus (Q', G', V') is an object of $\mathscr{C}_1$.

It is clear that this induces an isomorphism of the categories $\mathscr{C}$ and $\mathscr{C}_1$. This gives in the inverse limit :

Proposition 4 : Assume that $\mathscr{X}$ has a fundamental group-scheme at χ_0. Let $R = j^{-1}(\chi_1)$ in $j : P \to \mathscr{X}$. Then $\pi(\mathscr{X}, \chi_1)$ also exists, and

(a) R is a principal homogeneous space for both $\pi(\mathscr{X}, \chi_0)$ and $\pi(\mathscr{X}, \chi_1)$ and the actions of both on R commute, i.e.,

(b) $\pi(\mathscr{X}, \chi_1)$ is an inner twist of $\pi(\mathscr{X}, x_0)$; consequently,

(c) $\pi(\mathscr{X}, \chi_0)$ and $\pi(\mathscr{X}, \chi_1)$ are isomorphic after a base-change to $\bar{k}$, and

(d) $\pi(\mathscr{X}, \chi_0)_{ab}$ and $\pi(\mathscr{X}, \chi_1)_{ab}$ are isomorphic.

Now we examine the effect of base-change on fundamental group-schemes. It will be freely asumed that all schemes encountered have $\mathscr{P}$.

Let L be an arbitrary field extension of k. The base-change of a k-scheme Y to L will be denoted by $\bar{Y}$.

If $\mathscr{X}$ is a k-scheme and χ_0 a k-rational point, we can base-change the universal triple $(P, \pi(\mathscr{X}, \chi_0), \cdot)$ to L. If $(R, \pi(\bar{\mathscr{X}}, \bar{\chi}_0), \cdot)$ is the univesal triple for $(\bar{\mathscr{X}}, \bar{\chi}_0)$

by definition there is a unique morphism $(R, \pi(\bar{\mathcal{X}}, \bar{\chi}_0), \bar{*}) \to (\bar{P}, \overline{\pi(\mathcal{X}, \chi_0)}, \bar{*})$, and in particular, a natural homomorphism $\pi(\bar{\mathcal{X}}, \bar{\chi}_0) \to \overline{\pi(\mathcal{X}, \chi_0)}$.

If τ is an automorphism of L fixing k, again invoking the universal property, we see that there is a unique A_τ with a commutative diagram as below :

$$\begin{tikzcd} (R, \pi(\bar{X}, \bar{x}_0), \bar{*}) \arrow[r, "A_\tau"] \arrow[d, dash] & (R, \pi(\bar{X}, \bar{x}_0), \bar{*}) \arrow[d, dash] \\ \operatorname{Spec} L \arrow[r, "\tau"] & \operatorname{Spec} L \end{tikzcd}$$

It follows that $A_\sigma A_\tau = A_{\sigma\tau}$, and if L is a finite galois extension of k, this gives descent data for the above triple showing that there is a triple (Q, G, V) for the pair $(\mathcal{X}, \chi_0)$ such that $(\bar{Q}, \bar{G}, \bar{V}) = (R, \pi(\bar{\mathcal{X}}, \bar{\chi}_0), *)$. Using the unique morphism $(P, \pi(\mathcal{X}, \chi_0), *) \to (Q, G, V)$, we get a homomorphism $\pi(\mathcal{X}, \chi_0) \to G$ and therefore (after base-change) a homomorphism $\overline{\pi(\mathcal{X}, \chi_0)} \to \pi(\bar{\mathcal{X}}, \bar{x}_0)$. It is easy to see that these homomorphisms are inverses of each other : therefore $\pi(\bar{\mathcal{X}}, \bar{\chi}_0) \to \overline{\pi(\mathcal{X}, \chi_0)}$ is an isomorphisms.

If L is a finite separable extension of k, then there is a finite galois extension E if k containing L. Applying the above to E/k and E/L we see once again that $\pi(\bar{\mathcal{X}}, \bar{\chi}_0) \to \overline{\pi(\mathcal{X}, \chi_0)}$ is an isomorphism. This immediately gives

Proposition 5 : If L is an arbitrary separable algebraic extension of k, then $\pi(\bar{\mathcal{X}}, \bar{\chi}_0)$ and $\overline{\pi(\mathcal{X}, \chi_0)}$ are isomorphic.

Remark : If $\mathcal{X} = A^1$ and $L = k(t)$, then $\pi(\bar{\mathcal{X}}, \bar{\chi}_0)$ and $\overline{\pi(X, \chi_0)}$ are certainly not isomorphic (even if k is algebraically closed) as is easily seen by considering the Artin-coverings $Z - Z^p = f(T)$.

Here $L = \bar{K} =$ an algebraic closure of k. If $\mathcal{X} = A^1$ and k are not perfect, then $\pi(\bar{\mathcal{X}}, \bar{\chi}_0)$ and $\overline{\pi(\mathcal{X}, \chi_0)}$ are not isomorphic ; this is seen by comparing the α_p-quotients on either side.

However we believe that the following is true :

Conjecture : If $\mathcal{X}$ is complete, geometrically connected and reduced, and L is an arbitrary field extension of k, then $\pi(\bar{\mathcal{X}}, \bar{\chi}_0) \to \overline{\pi(\mathcal{X}, \chi_0)}$ is an isomorphism.

§ 2. All schemes considered are connected, reduced and of finite type over k, unless explicitly mentioned.

Proposition 6 : U is an open dense subset of $\mathcal{X}$ such that $\mathbf{O}_{X,x}$ is an integrally closed local domain for all $x \in \mathcal{X} - U$, and $f : Y \to \mathcal{X}$ is a morphism such that $\mathbf{O}_X \to f_* (\mathbf{O}_Y)$ when restricted to U is an isomorphism.

A. Let Q be a principal G-bundle on $\mathcal{X}$ (with G a finite group-scheme) such that the structure group of $f^* Q$ can be reduced to a closed subgroup-scheme H. Then the structure group of Q itself can be reduced to H.

B. If y_0 is a k-rational point of $f^{-1}(U)$ and $f(y_0) = x_0$, then $\pi(Y, y_0) \to \pi(\mathcal{X}, x_0)$ is a surjection.

Proof: Factor $j : Q \to X$ by $j' : Q \to Z = Q/H$ and $j'' : Z \to X$. By assumption, there is a $s : Y \to Z$ such that $j''. s = f$ with the principal H-bundle Q' on Y being $Y \times_Z Q$.

Now s induces $j''_*(\mathbf{O}_Z) \to f_*(\mathbf{O}_Y)$, and by restricting to U, $j''_*(\mathbf{O}_Z) \mid U \xrightarrow{\cong} \mathbf{O}_X \mid U \to f_*(\mathbf{O}_Y) \mid U$. Because j'' is an affine morphism, there is a $t : U \to Z$ such that $t \circ (f \mid f^{-1} U) = s \mid f^{-1} U$.

Let T be the closure of the image of t. Then, for all $x \in X - U$, $T \times_X \operatorname{spec} \mathbf{O}_{X,x} \to \operatorname{spec} \mathbf{O}_{X,x}$ is a finite birational morphism, and therefore an isomorphism. Therefore $T \to X$ is itself an isomorphism, thus giving a section $t : X \to Z$ such that $t \mid U = t$. Then $Q'' = j'^{-1}(t(X))$ is the required H-bundle on X. This finishes the proof of A.

To show that $\pi(Y, y_0) \to \pi(X, x_0)$ is a surjection, it suffices to show that the composite $\pi(Y, y_0) \to \pi(X, x_0) \to G$ is a surjection for all finite quotients G of $\pi(X, x_0)$.

Let G be one such, and let H be the image of $\pi(Y, y_0)$ in G. Then, there is a triple (Q, G, V) for the pair (X, x_0) and a triple (Q', H, V') for (Y, y_0) and a morphism $(Q', H, V') \to (f^*, Q, G, V\, xy_0)$.

By A there is a principal H-bundle Q'' on X and a diagram :

$$\begin{tikzcd}
Q' \mid f^{-1} U \arrow[r] \arrow[d] & Q' \arrow[d] \\
Q'' \mid U \arrow[r] & Q
\end{tikzcd}$$

thus showing that V is a k-rational point of $Q'' \mid U$. The inclusion (Q'', H, V), $\to (Q, G, v)$ shows that there is a factoring : $\pi(X, x_0) \to H \to G$. Consequently $H = G$ and B has also been proved.

Corollary: If $f : Y \to X$ is an open immersion and X is normal, then $\pi(Y, y_0) \to \pi(X, x_0)$ is a surjection.

Corollary : If $f : Y \to X$ is smooth and proper with connected fibres, then $\pi(Y, y_0) \to \pi(X, x_0)$ is a surjection.

Proposition 7 : Let U be open dense in X such that $\mathbf{O}_{X,x}$ is a regular local ring of dimension $\geqslant 2$, for all $x \in X - U$. Then any principal G-bundle on U (where G is a finite group-scheme) extends to one such on X.

If x_0 is a k-rational point of U, then $\pi(U, x_0) \to \pi(X, x_0)$ is an isomorphism.

Proof : The second assertion follows trivially from Proposition 6 and the first assertion.

Let $j : Q \to U$ be a principal G-bundle on U and $i : U \to X$ the inclusion morphism. Then $i_* j_*(\mathbf{O}_Q)$ is a coherent sheaf of $\mathbf{O}_X$-algebras and therefore there is a finite morphism $j' : P \to X$ such that $j'_*(\mathbf{O}_P) = i_* j_*(\mathbf{O}_Q)$. It is clear that there is a G-action on P such that j' is G-equivariant (with the trivial action of G on X), and that $j'^{-1}(U) = Q$.

Assume that $j'_*(\mathbf{O}_P)$ is locally free. Then the standard morphism $P \times G \to P \times_X P$ is an isomorphism if and only if it induces an isomorphism $j'_*(\mathbf{O}_P) \otimes j'_*(\mathbf{O}_P) \to j'_* \mathbf{O}_P \otimes R(G)$. But these are locally free and of the same rank and therefore if V is the largest open subset of X restricted to which it is an isomorphism, then

the complement of V is of pure codimension one. However, V contains U; consequently $V = X$ and P is a principal G-bundle on X.

Thus it suffices to show that $j'_*(\mathbf{O}_P)$ is locally free ; equivalently, $j' : P \to X$ is a flat morphism. For this purpose, we may clearly assume that $X = \operatorname{spec} R$, where R is a regular ring.

For a scheme Y in characteristic p, let $\pi_Y : Y \to Y$ be the Frobenius. If $Y =$ spec A, π_Y will also be noted by π_A and $(\pi_Y^m)_* \mathbf{O}_Y = \tilde{A}_m$. If A is an integral domain, $A_m = A^{1/p^m}$.

The commutative diagram

$$\begin{array}{ccc} G & \xrightarrow{\pi_G^m} & G \\ \downarrow & \pi_k^m & \downarrow \\ \operatorname{Spec} k & \longrightarrow & \operatorname{Spec} k \end{array}$$

induces a homomorphism $G \to (\pi_k^m)^* G$.

Similarly, the diagram

$$\begin{array}{ccc} Q & \xrightarrow{\pi_Q^m} & Q \\ \downarrow & \pi_U^m & \downarrow \\ U & \longrightarrow & U \end{array}$$

induces $Q \to (\pi_U^m)^* Q$. Also $(\pi_U^m)^* Q$ is easily seen to be a principal $(\pi_k^m)^* G$-bundle on U and the morphism $Q \to (\pi_U^m)^* Q$ intertwines the actions of G and $(\pi_k^m)^* G$.

The direct image of the structure sheaf under $(\pi_U^m)^* Q \to U \xrightarrow{\pi_U^m} U \to X$ is precisely $\widetilde{R_m \otimes_R B}$, where $\tilde{B} = i_* j_* (\mathbf{O}_Q)$, because R_m is R-flat. And the morphism $Q \to (\pi_U^m)^* Q$ induces the usual R-algebra homomorphism $R_m \otimes_R B \to B_m$ by taking direct images of structure sheaves.

Case 1 : If G is a local group-scheme, $G \to (\pi_k^m)^* G$ is the trivial homomorphism for a suitably large m. The morphism $Q \to (\pi_U^m)^* Q$ thus makes $(\pi_U^m)^* Q$ the trivial bundle on U. Therefore $R_m \otimes_R B$ is a free R_m-module, and because R_m is R-faithfully flat, B is a locally free R-module.

Case 2 : If G is geometrically reduced, then $G \to \pi_k^* G$ is an isomorphism ; it follows that $Q \to \pi_U^* Q$ and $R_1 \otimes_R B \to B_1$ are isomorphisms too. In particular, B_1 is B-flat. By a theorem of Kunze, this proves that B is regular. That B is R-flat follows from the fact that it is a finitely generated Cohen-Macaulay R-module,

In both cases we have shown that $j' : P \to X$ is flat.

Now for the general case : there is an exact sequence :

$1 \to G_{\text{loc}} \to G \to H \to 1$

where G_{loc} is a local group-scheme and H is geometrically reduced. In fact this sequence is split if k is assumed to be perfect.

In any case Q/G_{loc} is a principal H-bundle on U, and by case 2, Q/G_{loc} is the inverse image of U in a principal H-bundle $Z \to X$. The pair $(Z, Q/G_{loc})$ have the same properties as (X, U) and $Q \to Q/G_{loc}$ is a principal G_{loc}-bundle: by case 1, Q is the inverse image of Q/G_{loc} in a principal G_{loc}-bundle $W \to Z$. Thus Q is the inverse image of U in the flat morphism $W \to X$ showing that $i_* j_* (\mathbf{O}_Q)$ is locally free.

This completes the proof of Proposition 7.

Proposition 8: If X and Y are smooth complete birationally isomorphic varieties over k, and x_0, y_0 are geometric points of X and Y respectively, then $\pi(X, x_0)$ an $\pi(Y, y_0)$ are inner twists of each other.

If Y is assumed only to be normal and complete instead, and k is algebraically closed, then $\pi(Y, y_0)$ is a quotient of $\pi(X, x_0)$.

Proof: For the sake of simplicity, we shall prove the first statement only for separably closed field k. In this case, there is a complete normal variety Z, morphisms $f: Z \to X$ and $g: Z \to Y$, open subsets $U \to X$ and $V \to Y$ such that $f^{-1}(U) \to U$ and $g^{-1}(V) \to V$ are isomorphisms and a k-rational point z_0 in $f^{-1}(U) \cap g^{-1}(V)$. Put $x_0 = f(z_0)$ and $y_0 = g(z_0)$. Then we have :

$$\begin{tikzcd} \pi(f^{-1}(U), z_0) \arrow[r, "\text{onto}"] \arrow[d, "\cong"'] & \pi(Z, z_0) \arrow[d] \\ \pi(U, x_0) \arrow[r, "\cong"] & \pi(X, x_0) \end{tikzcd}$$

The above horizontal arrow is surjective by Proposition 6, and the one below is an isomorphism by Proposition 7. Therefore $\pi(Z, z_0) \to \pi(X, x_0)$ is an isomorphism. Similarly $\pi(Z, z_0) \to \pi(Y, y_0)$ is an isomorphism, and the general statement follows from Proposition 4.

If Y is only normal, then $\pi(Z, z_0) \to \pi(Y, y_0)$ is only a surjection, proving the second statement.

Lemma: If k is any field, $\pi(\mathbf{P}^1, x_0)$ is trivial.

Proof: The representations of $\pi(\mathbf{P}^1, x_0)$ are essentially finite bundles on $\mathbf{P}^1$. But any semi-stable bundle on $\mathbf{P}^1$ of degree zero is trivial, and therefore all representations of $\pi(\mathbf{P}^1, x_0)$ are trivial. Thus $\pi(\mathbf{P}^1, x_0)$ is itself trivial.

Proposition 9: Let $f: Z \to X$ be a smooth proper morphism with connected fibres. Assume that X is reduced. Also, for every t : spec $\bar{k} \to X$, the fibre Z_t has the trivial fundamentalgroup-scheme. Let z_0 be a k-rational point of Z and $f(z_0) = x_0$. Then $\pi(Z, z_0) \to \pi(X, x_0)$ is an isomorphism.

Proof: The surjectivity follows from the corollaries to Proposition 6. To prove the injectivity we have to show that any $j: P \to Z$ which is a principal G-bundle on Z is the pull-back of a principal G-bundle on X. Here G is a finite group-scheme.

For any t : spec $\bar{k} \to X$, $P \mid Z_t$ is the trivial bundle. Thus the semi-continuity theorem applied to the sheaf $j_*(\mathbf{O}_P)$ and the morphism $f: Z \to X$ shows that $f_* j_* (\mathbf{O}_P)$ is locally free, and, the natural homomorphism $f^* f_* j_* (\mathbf{O}_P) \to j_* (\mathbf{O}_P)$ is an isomorphism of $\mathbf{O}_Z$-algebras.

Thus, if $h : W \to X$ is defined by $h_* (O_W) = f_* j_* (O_P)$, i.e., $P \to W \xrightarrow{h} X$ is the Stein factorisation of $f \circ j$, then the natural map $P \to W \times_X Z$ is an isomorphism.

The Stein factorisation applied to the vertical arrows of

$$\begin{tikzcd} P \times G \arrow[rr] \arrow[dr] & & P \arrow[dl] \\ & X & \end{tikzcd}$$

gives

$$\begin{tikzcd} P \times G \arrow[rr] \arrow[d] & & P \arrow[d] \\ W \times G \arrow[rr] \arrow[dr] & & W \arrow[dl] \\ & X & \end{tikzcd}$$

Therefore there is a G-action on W, and $P \to W$ and $W \to X$ are G-equivariant (with the trivial action on X).

Also the standard $W \times G \to W \times_X W$ is an isomorphism because its base-change to Z is $P \times G \to P \times_Z P$ (which is by very assumption an isomorphism) and $Z \to X$ is flat and surjective.

This shows that W is a principal G-bundle on X and that $P = f^* W$, Q.E.D.

Corollary : Any complete normal rational variety has a trivial fundamental group-scheme.

Proof: By the above proposition, $\mathbf{P}^1 \times \mathbf{P}^1 \times \cdots \mathbf{P}^1$ has a trivial fundamental group-scheme. The corollary now follows from Proposition 8. No assumption on k is neessary; that k is algebraically closed was required in the proof of Proposition 8 only to get hold of a k-rational point in any non-empty open subset.

Corollary: If $f : Z \to X$ is smooth and proper with Z_t rational for all t : spec $\bar{k} \to X$, then $\pi (Z, z_0) \to (\pi (X, x_0)$ is an isomorphism.

This follows immediately from the above corollary and Proposition 9.

CHAPTER III

Parabolic bundles and ramified coverings

Let X be a smooth connected projective curve over an algebraically closed field k with a base-point x_0 and a finite subset S of X such that $x_0 \notin S$. We want to identify the representations of $\pi (X - S, x_0)$ with certain bundles on $X - S$ wih some additional structure (denoted by parabolic bundles) and show that the main theorems of Chapter I hold in this modified situation

First some generalities : denote by T the diagram

$$\begin{tikzcd} & Y_1 \\ Z \arrow[ur, "f_1"] \arrow[dr, "f_2"'] & \\ & Y_2 \end{tikzcd}$$

where Z, Y_1 and Y_2 are schemes over k and f_1 and f_2 are morphisms.

A vector bundle W on T is a $W = (V, V_1, V_2, \varphi_1, \varphi_2)$ where V, V_1 and V_2 are vector bundles on Z, Y_1 and Y_2 respectively and $\varphi_i : V \to f_i^* V_i$ are isomorphisms for $i = 1, 2$. It is clear what a homomorphism of vector bundles on T is, and what is meant by exactness of a sequence of vector bundles on T. The category of vector bundles on T shall be denoted by Vect T.

A principal G-bundle P on T is $P = (Q, Q_1, Q_2, \psi_1, \psi_2)$ where Q, Q_1 and Q_2 are principal G-bundles on Z, Y_1 and Y_2 respectively and $\psi_i : Q \to f_i^* Q_i$ are isomorphisms for $i = 1$ and 2. If Y_1 has a base point x_0, a base-point for P is a point of Q_1 sitting above x_0. Thus we can form the category of triples (P, G, v) where P is a principal G-bundle on T, G is a finite group-scheme, as in chapter II ; this category we shall denote by $\mathcal{C}(T)$.

An immediate extension of Proposition 2·9, Chapter I is :

Lemma 1 : There is a bijective correspondence between principal G-bundles P on T and functors $F : G\text{-mod} \to \text{Vect } T$ satisfying $F1$ to $F4$, where G is any affine group-scheme. The functor F associated to a principal G-bundle P on T will be denoted by $F(P)$ as usual.

In applications Z, Y_1 and Y_2 are going to be very special : $Y_1 = X - S$ with base-point x_0. For each $x \in S$, let K_x = quotient field of $\hat{O}_{X,x}$ and E_x an arbitrary algebraic extension of K_x. Let R_x be the integral closure of $\hat{O}_{X,x}$ in E_x. Put $Z = \operatorname{spec}(\bigoplus_{x \in S} E_x)$ and $Y_2 = \operatorname{spec}(\bigoplus_{x \in S} R_x)$, with the obvious morphisms f_1 and f_2.

With this choice of T, a parabolic bundle on $X - S$ is just a vector bundle on T, and homomorphisms of parabolic bundles are just vector bundles homomorphisms on T.

2 : For each $x \in S$ choose an isomorphism of K_x with $k(t)$ and put $E_x = \bigcup_{n \geq 1} k(t^{1/n})$. Then parabolic bundles on $X - S$ are precisely "parabolic bundles with fractional weights" in the sense of Seshadri.

3 : However, most frequently we shall put $E_x = \bar{K}_x$, the algebraic closure of K_x. One good reason is the following :

Lemma 2 : $\mathcal{C}(X - S, x_0)$ be the category of triples (Q, G, v) associated to the pair $(X - S, x_0)$. If $E_x = \bar{K}_x$ for all $x \in S$, then $\mathcal{C}(T) \to \mathcal{C}(X - S, x_0)$ is an isomorphism.

Proof : A principal G-bundle on spec E_x can be regarded as a principal homogeneous space for spec $E_x \times_{\operatorname{spec} k} G$, but E_x being algebraically closed, this always admits a E_x-rational point. Thus any principal G-bundle on spec E_x is trivial. By Proposition 6, Chapter II, it follows that any principal G-bundle on spec R_x is trivial, where G is a finite k-group-scheme.

Let P be a principal G-bundle on T. Then $P = (Q, Q_1, Q_2, \psi_1, \psi_2)$, and we may assume that $Q = Z \times G$, $Q_2 = Y_2 \times G$ and ψ_2 is the obvious isomorphism from Q to $f_2^* Q_2$. Thus, such a P is completely determined by a principal G-bundle Q_1 on $Y_1 = X - S$, and an isomorphism $\psi_1 : Z \times G \to f_1^* Q_1$.

It follows that the objects of $\mathcal{C}(T)$ can be identified with (Q_1, G, ψ_1, v) where Q_1 is principal G-bundle on $X - S$, G a finite group-scheme, v a base-point of Q_1 above x_0, and $\psi_1 : Z \times G \xrightarrow{\cong} f_1^* Q_1$.

Given (Q_1, G, v) an object of $\mathcal{C}(X - S, x_0)$, such a ψ_1 always exists because all G-bundles on Z are trivial. This shows that $\mathcal{C}(T) \to \mathcal{C}(X - S, x_0)$ is a "surjection".

A morphism from (Q_1', G', ψ_1', v') to $(Q_1'', G'', \psi_1'', v'')$ in $\mathcal{C}(T)$ is by definition a (h, h_1, h_2, ρ) such that $A : (h_1, \rho) : (Q_1', G', v') \to (Q_1'', G'', v'')$ is a morphism in $\mathcal{C}(X - S, x_0)$ and a commutative diagram

B
$$\begin{tikzcd}
Y_2 \times G' \arrow[d, "h_2"'] & Z \times G' \arrow[l] \arrow[r] \arrow[d, "h"] & f_1^* Q_1' \arrow[d, "f_1^* h_1"] \\
Y_2 \times G'' & Z \times G'' \arrow[l] \arrow[r] & f_1^* Q_1''
\end{tikzcd}$$

where h and h_2 intertwine the right actions of G' and G'' *via* $\rho : G' \to G''$.

Given a (h_1, ρ) satisfying A, we shall show that there is a unique (h, h_1, h_2, ρ) satisfying both A and B. The right side of the diagram determines h. Now h is equivalent to giving a morphism $Z \to G''$ and this necessarily factors : $Z \to \operatorname{spec} k \xrightarrow{l} G''$. Using the morphism $Y_2 \xrightarrow{l} \operatorname{spec} k \to G''$, one gets the required h_2.

This is exactly the same as saying that $\mathcal{C}(T) \to \mathcal{C}(X - S, x_0)$ is fully faithful. The lemma is now proved.

We now define the degree of a parabolic vector bundle.

Let $v_s : E_s \to Q \cup \infty$ be the unique valuation such that $v_s(K_s) = \mathbf{Z} \cup \infty$. If L is a line bundle on Y_2 and s is a section of $f_2^* L$ on Z, then $(\bigoplus_{s \in S} R_s)\, s = \bigoplus_{s \in S} h_s L_s$ for some $h_s \in E_s$, where L_s is the stalk of the sheaf of sections of L at x. Define $v_s(s)$ to be $v_s(h_s)$.

If $W = (L, L_1, L_2)$ is a line bundle on T, and s is a section of L_1 on $Y_1 = X - S$, put $\deg s = \sum_{x \in X} v_x(s)$. For $x \in X - S$, $v_x(s)$ is as usual the order of vanshing of s at x, and if $x \in S$, then $v_x(s)$ makes sense as above, as a rational number. Then deg s is easily seen to be independent of the choice of s and we define $\deg W = \deg s$.

If W is a parabolic vector bundle of rank r, then we define $\deg W = \deg \Lambda^r W$. Clearly we have :

Lemma 3 : A. For an exact sequence $0 \to V' \to V \to V'' \to 0$ of parabolic bundles, $\deg V' + \deg V'' = \deg V$.

B. If $f : V \to W$ is a homomorphism of parabolic bundles and $f \mid U$ is an isomorphism for some open subset U of $X - S$, then $\deg V \leqslant \deg W$.

C. If $\mu(V) = \deg V / rk\, V$, then $\mu(V \otimes W) = \mu(V) + \mu(W)$.

A homomorphism $h: W \to V$ of parabolic bundles is a generic injection if $h|U$ is an injection for some nonempty open subset U of $X - S$.

A parabolic bundle V of degree zero is semi-stable if for all $h: W \to V$ which are generic injections, $\deg W \leqslant 0$.

Lemma 4 : A. Parabolic semi-stable bundles of degree zero form an abelian category.

B. If V is a parabolic bundle of rank $r > 1$, there is an exact sequence $0 \to V' \to V \to V'' \to 0$ with $rkV' = 1$.

C. If $0 \to V' \to V \to V'' \to 0$ is an exact sequence of parabolic bundles and $h: W \to V$ is a generic injection, there is a diagram :

$$\begin{array}{ccccccccc} 0 & \to & W' & \to & W & \to & W'' & \to & 0 \\ & & \downarrow{\scriptstyle h'} & & \downarrow{\scriptstyle h} & & \downarrow{\scriptstyle h''} & & \\ 0 & \to & V' & \to & V & \to & V'' & \to & 0 \end{array}$$

with the horizontal arrows exact, and h' and h'' are generic injections.

The proofs are simple enough. For example C is proved by showing that such a diagram exists on Z, Y_1 and Y_2 separately (because torsion-free modules are locally free) and patching up.

A finite bundle is a parabolic bundle V such that $f(V) \cong g(V)$ for some $f \neq g$ which are polynomials with non-negative integer coefficients. An essentially finite bundle is defined exactly as in Chapter I.

Lemma 5 : The global sections of a parabolic bundle are finite dimensional.

Proof : Let (V, V_1, V_2) be a parabolic bundle. Let W be any vector bundle on X such that $W \mid X - S = V_1$. Then it is easy to see that there is a N such that the global sections of this parabolic bundle are contained in $\Gamma(X, W(ND))$ where $D = \sum_{x \in S} x$. This proves the lemma.

Lemmas 3, 4 and 5 immediately show that all the results of § 3, Chapter I up to Proposition 3·7 are valid for finite and essentially finite parabolic bundles.

Proposition 1 : Let P be a principal G-bundle on T, and G a finite group-scheme. Then, for all representations W of G, $F(P)\,W$ is an essentially finite parabolic bundle.

Proof : Let R be the co-ordinate ring of G, and if $n = rkR$, then $R \otimes R = R^n$ as G-representations. Consequently, if $V = F(P)\,R$, then $V \otimes V = V^n$, i.e., $[V]$ satisfies the polynomial $x^2 = nx$.

If W is a representation of G, then $W \to R^m$ for some m, and therefore $F(P)W \to V^m$. It suffices to show that $F(P)\,W$ has degree zero to prove that it is essentially finite. But if $r = rkW$, $L = \Lambda^r W$, $E = F(P)\,L$, then $\Lambda^r(F(P)\,W) = E$ and E is a parabolic line bundle of finite order and therefore has degree zero.

Exactly as in Chapter I, we see that all essentially finite bundles on T form a Tannaka category. Let $\pi(T)$ be the associated group-scheme. Then

Proposition 2 : Objects of $\mathcal{C}(T)$ are in one-to-one correspondence with homomorphisms from $\pi(T) \to G$.

Combining this with Lemma 2, we have :

Proposition 3 : If $E_x = \bar{K}_x$ for all $x \in S$, $\pi(T) = \pi(X - S, x_0)$ and its category of representations is precisely all essentially finite parabolic bundles on $X - S$.

If Char $k = 0$, note that $E_x = \bar{K}_x$ is the same as example 2, so that we have a good description of parabolic bundles in this case.

PART II

CHAPTER IV

1. Nilpotent bundles and formal groups for curves

An affine group-scheme G is nilpotent if every G-representation $V \neq 0$ has a $v \neq 0$ which is fixed by G.

Let X be a k-scheme of finite type with $\Gamma(X, \mathbf{O}_X) = k$. We shall be concerned with principal G-burdles P on X.

Lemma 1 : If P and G are as above, and V is a finite dimensional representation of G, then $W = F(P)\, V$ has a filtration $W = W_0 \supseteq W_1 \supseteq \cdots \supseteq W_n = 0$ such that W_i / W_{i+1} is a trivial vector bundle on X.

Proof : By induction, we see that V has a filtration $V = V_0 \supseteq V_1 \supseteq \cdots \supseteq V_n = 0$ such that V_i / V_{i+1} is a trivial representation of G. Simply put $W_j = F(P)\, V_j$.

The full subcategory of all vector bundles on X with objects as those vector bundles on X that admit a filtration as in Lemma 1 will be denoted by $N(X)$.

Lemma 2 : $N(X)$ is an abelian category and is closed with respect to tensor products and duals.

Proof : Let $f : V \to W$ be a homomorphism with V and W in $N(X)$. For convenience, we identify vector bundles with their sheaves of sections. We have to show that Ker (f) and Coker (f) are locally free, and moreover, belong to $N(X)$. We shall do so by induction on $rk\, V + rk\, W$.

If $V = 0$ or $W = 0$, or if $rk\, V = rk\, W = 1$, then the statement is obvious because $\Gamma(X, \mathbf{O}_X) = k$.

Also, the statement for f is equivalent for the statement for $f^* : W^* \to V^*$. Therefore, replacing f by f^* if necessary, we may assume that $rk\, W > 1$. There is an exact sequence $0 \xrightarrow{i} W' \xrightarrow{j} W \to \mathbf{O} \to 0$ with W' in $N(X)$. Also $rk\, V + rk\, \mathbf{O}_x < rk\, V + rk\, W$, so the statement holds for $j \circ f$. Thus $j \circ f = 0$ or $j \circ f$ is an isomorphism.

In the first case, we have $g : V \to W'$ such that $i \circ g = f$, and by the induction hypothesis ker (g) and coker (g) are locally free and in $N(X)$. But clearly, ker (f) = ker (g) and $0 \to$ coker $(g) \to$ coker $(f) \to \mathbf{O}_x \to 0$ is exact, so that ker (f) and coker (f) have the required property.

The category of modules over a ring is denoted both by A-mod and $| A |$ in the text,

In the second case, $V' = f^{-1}(W') = \ker(j \circ f)$ is in $N(X)$ and $h : V' \to W$ defined by

$$\begin{tikzcd} V' \arrow[r, "h"] \arrow[d] & W' \arrow[d] \\ V \arrow[r, "f"] & W \end{tikzcd}$$

has the required property by the induction hypothesis, and here h and f have the same kernels and cokernels.

That $V \otimes W$ is in $N(X)$ if both V and W are in $N(X)$ is obvious.

Let x_0 be a k-rational point of X and let $T : N(X) \to k$-mod be given by $T(V)$ = fibre of V at x_0. Then $N(X)$ becomes a Tannaka category and by Chapter I, § 1, there is an affine group-scheme $U(X, x_0)$ such that $U(X, x_0)$-mod is isomorphic to $N(X)$.

Every non-zero vector bundle in $N(X)$ has a trivial sub-bundle; consequently every non-zero representation of $U(X, x_0)$ contains a non-zero fixed subspace. In other words $U(X, x_0)$ is a nilpotent group-scheme. By § 2, Chapter I, we see that there is a principal $U(X, x_0)$-bundle P on X with a k-rational point $*$ above x_0 with the following universal property:

Proposition 1: For every principal G-bundle Q on X with a k-rational point v of Q above x_0 (and G nilpotent), there is a unique homomorphism $\rho : U(X, x_0) \to G, f : P \to Q$ intertwining the actions of $U(X, x_0)$ and G, such that $f(*) = ve$.

Proposition 2: $\mathrm{Hom}(U(X, x_0), G_a) = H^1(X, \mathrm{O}_X)$. In particular, if characteristic $k = 0$, then $U(X, x_0)_{ab} = H^1(X, \mathrm{O}_X)^*$.

Proof: The second assertion follows immediately from the first.

The first follows from the well-known fact: isomorphism classes of principal G_a-bundles are in one-to-one correspondence with members of $H^1(X, \mathrm{O}_X)$.

We recall some basic facts about affine group-schemes: if $\mu : R \to R \otimes R$ is the co-multiplication of R, the coordinate ring of $U(X, x_0)$, and if $A = R^* =$ the vector space dual of R, then A is an algebra. The collection of $V^{\perp}$, where the V range through all finite dimensional subspaces of R such that $\mu V \subseteq V \otimes V$, give a system of neighdourhoods of zero under which A is complete: $A = \varprojlim_{V} A/V^{\perp}$. Note that $A/V^{\perp}$ is a finite-dimensional k-algebra.

In fact any open two-sided ideal is of the form $V^{\perp}$ where V is of the above type.

Also, $U(X, x_0)$-mod is the category A-mod of left A-modules M such that each member of M is annihilated by some V. Because $U(X, x_0)$ is nilpotent, it follows that each $A/V^{\perp}$ is an Artin-local ring (with maximal ideal $\mathbf{m}/V^{\perp}$ where $\mathbf{m} = k^{\perp}$ and $A/\mathbf{m} = k$).

Let $J_n = \bigcap_V (\mathbf{m}^n + V^{\perp})$, i.e., J_n is the closure of $\mathbf{m}^n$ in A.

Lemma 3: Assume that $H^1(X, \mathrm{O})_X$ is finite dimensional. Then

(a) J_n is open for all n,

(b) Every open two-sided ideal contains J_n for some n. Therefore $A = \varprojlim_n A/J_n$.

(c) $(J_1/J_2)^* = H^1(\mathfrak{X}, O_X)$.

(d) $J_n = \mathbf{m}^n$ for all n.

Proof ; We shall have no occasion to use D, so we omit its proof. Let J be any open two-sided ideal. Let $\bar{\mathbf{m}} = \mathbf{m}/J$ be the maximal ideal of the Artin-local ring A/J. Then $(\bar{m}/\bar{m}^2)^* = (m/m^2 + J)^* = \mathrm{Der}\,(A/J, k) =$ all derivations $D : A \to k$ that vanish on J. If $J = W^{\perp}$, all such derivations are in 1–1 correspondence with $x \in W$ such that $\mu \mathfrak{X} = X \otimes 1 + 1 \otimes \mathfrak{X}$. This in turn is a subspace of $\mathrm{Hom}\,(U(\mathfrak{X}, x_0), G_a) = L = \{x \in R \mid \mu\mathfrak{X} = \mathfrak{X} \otimes 1 + 1 \otimes x\} = H^1(\mathfrak{X}, O_X)$. Let $rk\, L = g$. By assumption, g is finite. Then

$$rk\,(A/\mathbf{m}^n + J) \leqslant 1 + rk\,(\bar{\mathbf{m}}^2/\bar{\mathbf{m}}^3) + rk\,(\bar{\mathbf{m}}^2/\bar{\mathbf{m}}^3) + \cdots + rk + (\bar{\mathbf{m}}^{n-1}/\bar{\mathbf{m}}^n)$$
$$\leqslant 1 + g + g^2 + \cdots + g^{n-1}.$$

From this it follows that there is an open two-sided ideal E such that for all open two-sided J contained in E, $A/\mathbf{m}^n + J \to A/\mathbf{m}^n + E$ is an isomorphism. This shows clearly that $J_n = \mathbf{m}^n + E$. This proves (a).

By the above remarks, if $J = W^{\perp}$, then $J + \mathbf{m}^2 = (k + W \cap L)^{\perp}$. It follows that $J_2 = (k + L)^{\perp}$. Thus $(J_1/J_2)^* = k + L/k = L$. This proves (c).

If J is any open two-sided ideal, then some power of the maximal ideal of A/J is zero. In other words, J contains some power of the maximal ideal, and therefore contains its closure, which is J_n for some n. This proves (b).

Let $R_n \subseteq R$ be defined by $R_n^{\perp} = J_n$. Then

Lemma 4 :

(a) $\mu(R_n) \subseteq R_n \otimes R_n$.

(b) R_n is finite dimensional.

(c) the R_n span R.

Proof : (a) and (b) follow from the fact that J_n is an open two-sided ideal (c) follows from dualising (b) of the above lemma.

We shall now get an explicit definition of the R_n. Let $R^{\otimes n}$ be the n-fold tensor product of R and $\mu_n : R \to R^{\otimes n}$ the iterated co-multiplication map. In fact it is induced by the multiplication morphism $G \times G \times \cdots \times G \to G$ where $G = U(\mathfrak{X}, x_0)$. Let i_t: $R^{\otimes (n-1)} \to R^{\otimes n}$ be the inclusion by tensoring with 1 in the t-th factor, for $1 \geqslant t \geqslant n$. Let $S_n =$ span of the images of $i_1, i_2, \cdots, i_n$.

Clearly, $(R^{\otimes n})^* = A \hat{\otimes} A \hat{\otimes} \cdots \hat{\otimes} A$ where $\hat{\otimes}$ denotes the completed tensor product, and $S_n = (m \hat{\otimes} m \hat{\otimes} \cdots \hat{\otimes} m)^{\perp}$. Then $\bar{J}_n$ is the closure of the image of S_n in $A \hat{\otimes} A \hat{\otimes} \cdots \hat{\otimes} A \to A$, and therefore $J_n = R_n^{\perp}$, where $R_n = \{x \in R \mid \mu_n x \in S_n\}$.

The following completely distinguishes the situation in characteristic zero and positive characteristic (assuming as always that $rk\, H^1(X, \mathbf{O}_X) < \infty$).

Proposition 3 : If characteristic $k = p > 0$, then $U(X, x_0)$ is an inverse limit of finite group-schemes.

If X has a fundamental group-scheme, by the universal properties enjoyed by both $\pi(X, x_0)$ and $U(X, x_0)$, it follows that $U(X, x_0)$ is a quotient of $\pi(X, x_0)$.

Proof : Let H_n be the k-subalgebra of R generated by R_n. By lemma 4 (a), H_n is a Hopf sub-algebra of R.

If $\theta \in R_n$, then $\mu_n \theta \in S_n$. We have : $\mu_n(\theta^p) = (\mu_n \theta)^p \in S_n^p \subseteq S_n$. Therefore θ^p also belongs to R_n.

Let $x_1, x_2, \cdots, x_l$ span R_n. Then $x_1^{a_1} x_2^{a_2} \cdots x_l^{a_l}$ for $0 \leqslant a_i \leqslant p-1$ span H_n. because R_n is closed with respect to p-th powers. Therefore H_n is finite dimensional.

By lemma 4 (c), R is the union of its finite dimensional Hopf subalgebras, i.e., $U(X, x_0)$ is an inverse limit of finite group-schemes.

For simplicity, we assume that X is complete from now on.

Definition: An exact sequence $0 \to \mathbf{O}_X^m \to U(V) \to V \to 0$ is a universal extension of V, if for every exact sequence $0 \to \mathbf{O}_X^n \to W' \to V \to 0$ there is a diagram :

$$\begin{tikzcd} 0 \arrow[r] & \mathbf{O}_X^m \arrow[r] \arrow[d, "f"'] & U(V) \arrow[r] \arrow[d] & V \arrow[r] \arrow[d, "1"] & 0 \\ 0 \arrow[r] & \mathbf{O}_X^n \arrow[r] & W' \arrow[r] & V \arrow[r] & 0 \end{tikzcd}$$

and while this diagram is not necessarily unique, the $f : \mathbf{O}_X^m \to \mathbf{O}_X^n$ is unique.

Every vector bundle V on X has a universal extension : Choose a basis $\xi_1, \xi_2, \cdots, \xi_m$ of $H^1(X, V^*)$. Then $\theta = (\xi_1, \xi_2, \cdots, \xi_m) \in H^1(X, V^*) = H^1(X, \mathrm{Hom}(V, \mathbf{O}_X^m))$ defines the required extension :

$$0 \to \mathbf{O}_X^m \to U(V) \to V \to 0.$$

More canonically, the $\mathbf{O}_X^m$ should be replaced by $j^*(H^1(X, V^*))^*$ where $j : X \to \mathrm{Spec}\, k$.

An immediate consequence of the definition is :

Lemma 5 : The natural map $H^1(X, V^*) \to H^1(X, U(V)^*)$ is identically zero,

Similarly, for A-modulus M, a universal extension of M is an exact sequence : $0 \to k^m \to U(M) \to M \to 0$ of A-modules with an identical universal property.

Denote by F the natural equivalence from $|U(X, x_0)| = |A| \to N(X)$. The next lemma is obvious.

Lemma 6 : $U(FM) = F(U(M))$ for all A-modules M.

Lemma 7 : If $V_1 = \mathbf{O}_X$ and $V_{n+1} = U(V_n)$ for all n, then $F(A/J_n) = V_n$ for all n.

Proof : By Lemma 6, it suffices to show that $U(A/J_n) = A/J_{n+1}$ for all $n \geqslant 1$.

Let $0 \to k^m \to M \to A/J_n \to 0$ be an exact sequence of A-modules annihilated by some open two-sided ideal E. But M is clearly annihilated by m^{n+1} ; there-

fore its annihilator contains $m^{n+1} + E \supset J_{n+1}$. Commutative diagrams below are clearly in one-to-one correspondence with elements $\xi \in M$ which go to $\bar{1}$ in A/J_n :

$$\begin{array}{ccccccccc} 0 & \longrightarrow & J_n/J_{n+1} & \longrightarrow & A/J_{n+1} & \longrightarrow & A/J_n & \longrightarrow & 0 \\ & & \downarrow h & & \downarrow & & \downarrow 1 & & \\ & & k^m & \longrightarrow & M & \longrightarrow & A/J_n & \longrightarrow & 0 \end{array}$$

But any two choices of ξ differ by an element of k^m which is annihilated by J_1, showing that h does not depend on the choice of ξ. Q.E.D.

Proposition 4 : If $\dim X = 1$, then $A \cong k\{\{X_1, X_2, \cdots, X\}\}$ which is the non-commutative formal power series ring in g variables, and $g = rk\, H^1(X, O_X)$.

Proof : $A = \varprojlim_n A/J_n$. Choose $x_1, x_2, \cdots, x_g \in J_1$ so that they form a basis for J_1/J_2. Putting $B = k\{\{X_1, X_2 \cdots, X_g\}\}$ and sending X_j to the x_j for all j, there is a natural surjection $B \to A$. This is an isomorphism if and only if $B/m^n \to A/J_n$ is an isomorphism for all n. For this it suffices to show that

$$rk\,(A/J_n) = rk\,(B/m^n) = 1 + g + g^2 + \cdots + g^{n-1}.$$

But $rk\,(A/J_n) = rk\,(V_n)$ where the V_n are as in Lemma 7.

If $0 \to O_X^l \to U(V_n) = V_{n+1} \to V_n \to 0$ is the universal extension of V_n, then $H^1(X, V_n^*) \xrightarrow{0} H^1(X, V_{n+1}^*) \to H^1(X, O_X^l) \to H^2(X, V_n^*) = 0$ is exact and therefore $rk\, H^1(X, V_{n+1}^*) = lg$

$$= g\, rk\, H^1(X, V_n^*).$$

Therefore $rk\, H^1(X, V_n^*) = g^n$ for all n and $rk\, V_n = 1 + g + \cdots + g^{n-1}$ for all n. This proves the Proposition.

We shall use this Proposition a little later.

We know what $U(X, x_0)_{ab}$ is in characteristic zero. We attempt below to understand this in the general case.

First we need :

Definition : Pic X is the following functor from k-schemes to abelian groups for a k-scheme S, a member of Mor $(S, \text{Pic}\, X)$ is a line bundle on $X \times S$ with a chosen trivialisation on $x_0 \times S$.

Proposition 5 : Let B be a commutative local finite-dimensional (Artin) k-algebra with $B/m = k$.

The following data are equivalent :

1. A k-algebra homomorphism $A \to B$ that vanishes on some J_n ;
2. An element of Mor (Spec B, Pic X) which vanishes when restricted to Mor (Spec k, Pic X).

Proof : We are given a line bundle L on $Y = X \times \text{Spec}\, B$ with a trivialisation on $x_0 \times \text{Spec}\, B$ and $X \times \text{Spec}\, k$. Let p_1 and p_2 be the projections. For a B-module M, let $GM = (p_1)_* (L \otimes p_2^* \tilde{M})$.

This is a functor from $|B|$ to $N(X)$ such that there is a natural equivalence of the functors $T \circ G$ and the forgetful functor from $|B|$ to $|A|$. Recall that $T : N(X) \to |k|$ is defined by $T(V) =$ fibre of V at x_0. This gives

$$\begin{tikzcd} B\text{-mod} \arrow[r, "G"] \arrow[dr] & N(X) \arrow[d, "T"] & A\text{-mod} \arrow[l] \arrow[dl] \\ & k\text{-mod} & \end{tikzcd}$$

and therefore a functor from $|B|$ to $|A|$ which respects the forgetful functors from both to $|k|$. Applying this functor to B itself, we see that B becomes an A-module such that $R_\alpha =$ right multiplication by α for $\alpha \in B$ is a A-module homomorphism for all $\alpha \in B$. Consequently there is a k-algebra homomorphism from A to B.

Conversely, given $j : A \to B$, define G to be the composite : $|B| \to |A| \xrightarrow{F} N(X)$. Because $B = \mathrm{End}_{|B|}(B, B)$, it follows that $G(B)$ has an action of B; equivalently, $G(B) = (p_1)_* L$ where L is a sheaf on $Y = X \times \operatorname{Spec} B$. We omit the checking that L is an invertible sheaf on Y.

As a consequence, we have :

Proposition 6 : Assume Pic X is representable. Then

1. $A_{ab} =$ the completion of the local ring of Pic X at zero.

2. The natural homomorphism $A_{ab} \to A_{ab} \hat{\otimes} A_{ab}$ is induced by taking completions at zero of Pic $X \times$ Pic $X \to$ Pic X.

3. If char $k = p > 0$, then $U(X, x_0)_{ab} = \varprojlim_G \hat{G}$ where the inverse limit is taken over all local group-schemes G embedded in Pic X.

Proof ; 1 and 2 follows immediately from the previous Proposition. For 3, observe that

(a) The dual of the co-ordinate ring of $U(X, x_0)_{ab}$ is just A_{ab}.

(b) if $G_n \hookrightarrow$ Pic X is the local group-scheme defined by the ideal J_n generated by p^nth powers of all elements of the maximal ideal, then $A_{ab} = \varprojlim_n R(G_n)$, where $R(G_n)$ is the co-ordinate ring of G_n, and

(c) $R(\hat{G}_n) = R(G_n)^*$.

A morphism $(Y, y_0) \to (X, x_0)$ clearly induces a homomorphism $U(Y, y_0) \to U(X, x_0)$. The projections $X \times Y \to X$ and $X \times Y \to Y$ thus induce a homomorphism

$$U(X \times Y, x_0 \times y_0) \to U(X, x_0) \times U(Y, y_0).$$

Lemma 8 : The above homomorphism is an isomorphism.

Corollary : If X is an abelian variety, then $U(X, 0)$ is abelian.

Proof: The multiplication $X \times X \to X$ clearly induces $U(X \times X, 0 \times 0) = U(X, 0) \times U(X, 0) \to U(X, 0)$ which is just the multiplication in the group-scheme $(U(X, x_0)$. Because this is a homomorphism, $U(X, 0)$ is commutative.

Proof of Lemma: $i : X \to X \times Y$ defined by $X \to X \times y_0 \to X \times Y$ induces $U(X, x_0) \to U(X \times Y, x_0 \times y_0)$. Because $p_1 \circ i = 1_X$ and $p_2 \circ i =$ constant, it follows that the composite $U(X, x_0) \to U(X \times Y, x_0 \times y_0) \to U(X, x_0) \times U(Y, y_0)$ is just $\alpha \mapsto (\alpha, 0)$.

Thus we see that $U(X \times Y, x_0 \times y_0) \to U(X, x_0) \times U(Y, y_0)$ is surjective. To show that it is injective also, it suffices to show that any representation V of $U(X \times Y, x_0 \times y_0)$ is a quotient of $P \otimes Q$ where P and Q are representations of $U(X, x_0)$ and $U(Y, y_0)$ respectively. Or, what is the same, to show that any V in $N(X \times Y)$ is a quotient of $P \otimes Q$ where P and Q are in $N(X)$ and $N(Y)$ respectively.

Sublemma: If V and W are in $N(X)$ and $N(Y)$ respectively, then $U(V \otimes W)$ is a quotient of $U(V) \otimes U(W)$.

Proof: Let $0 \to H^1(X, V^*)^* \otimes_k \mathbf{O}_X \to U(V) \to V \to 0$ and $0 \to H^1(Y, W^*)^* \otimes_k \mathbf{O}_Y \to U(W) \to W \to 0$ be the universal extensions. The quotient of $U(V) \otimes U(W)$ by $H^1(X, V^*)^* \otimes H^1(Y, W^*)^* \otimes \mathbf{O}_{X\times Y}$ gives an exact sequence :

$$0 \to H^1(X, V^*)^* \otimes_k \mathbf{O}_X \otimes W \oplus H^1(Y, W^*)^* \otimes_k V \otimes \mathbf{O}_Y \to Z \to V \otimes W \to 0.$$

Also there are canonical surjections $V \to H^\circ(X, V^*)^* \otimes_k \mathbf{O}_X$ and $W \to H^\circ(Y, W^*)^* \otimes_k \mathbf{O}_Y$. This gives a surjection from Z to Z' with :

$$0 \to (H^1(X, V^*)^* \otimes H^\circ(Y, W^*)^* \oplus H^1(Y, W^*)^* \otimes H^\circ(X, V^*)^*) \otimes \mathbf{O}_{X\times Y} \to Z' \to V \otimes W \to 0$$

which by the Kunneth formula is easily seen to be the universal extension of $V \otimes W$. Q.E.D.

Let V_n be as in Lemma 7 and let W_n be the same sequence for Y. Using this sublemma, we shall show that any T in $N(X \times Y)$ of rank n is a quotient of $(V_n \otimes W_n)^n$ by induction on n.

If $n = 1$, this is obvious.

If $rk\,T = n + 1$, $0 \to \mathbf{O}_{X\times Y} \to T \to T' \to 0$ is exact with T' in $N(X \times Y)$. There is a surjection $(V_n \otimes W_n)^n \to T'$ by which one pulls back this extension by $\mathbf{O}_{X\times Y}$ to get the following diagram :

$$\begin{tikzcd}
 & & (V_{n+1} \otimes W_{n+1})^n \arrow[r] \arrow[d] & (V_n \otimes W_n)^n \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & \mathbf{O}^{l}_{X\times Y} \arrow[r] \arrow[d] & U(V_n \otimes W_n)^n \arrow[r] \arrow[d] & (V_n \otimes W_n)^n \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & \mathbf{O}_{X\times Y} \arrow[r] \arrow[d, "1"] & H \arrow[r] \arrow[d] & (V_n \; W_n)^n \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & \mathbf{O}_{X\times Y} \arrow[r] & T \arrow[r] & T' \arrow[r] & 0
\end{tikzcd}$$

If $(V_{n+1} \otimes W_{n+1})^n \to T$ is not surjective, its image is a sub-bundle N of T such that $\mathbf{O}_{X\times Y} \oplus N = T$. In either case, it is clear that T is a quotient of $(V_{n+1} \otimes W_{n+1})^{n+1}$.

This finishes the proof of Lemma 8.

We come back to curves. Fix a $g \geqslant 1$.

We denote by $k\{\{X;Y;Z;\cdots\}\}$ the non-commutative formal power series ring in $X_1, X_2, \cdots, X_g, Y_1, Y_2, \cdots, Y_g, Z_1, Z_2, \cdots Z_g, \cdots$, modulo the relations : $X_i \cdot Y_j = Y_j \cdot X_i$ for all i and j, $X_i \cdot Z_j = Z_j \cdot X_i$ for all i and j, $Y_i \cdot Z_j = Z_j \cdot Y_i$ for all i and j, etc.

A non-commutative formal group is a $F(X;Y) = (F_1(X;Y), F_2(X;Y), \cdots F_g(X;Y))$ with the $F_i(X,Y) \in k\{\{X;Y\}\}$ such that

(a) $F_i(X,0) = F_i(0;X) = X_i$ for $1 \leqslant i \leqslant g$.

(b) $F(X;Y) = F(Y;X)$.

(c) $F(F(X:Y);Z) = F(X;F(Y;Z))$. This is an identity in $k\{\{X;Y;Z\}\}$.

Every complete curve X with $\Gamma(X, \mathbf{O}_X) = k$ gives rise to such a formal group: with A as usual, there is an isomorphism $A \cong k\{\{X\}\}$ and the homomorphism $k\{\{X\}\} \to A \cong A \hat{\otimes} A \cong k\{\{X;Y\}\}$ provides the $F_i(X;Y)$: put $F_i(X;Y) =$ the image of X_i.

Given a non-commutative formal group, put R = continuous linear functionals on $k\{\{X\}\}$. Then R is a Hopf algebra and G = Spec R is an affine nilpotent group-scheme with $R^* = k\{\{X\}\}$.

In characteristic zero, after a change of co-ordinates, $F_i(X;Y) = X_i + Y_i$. The lie algebra of the affine nilpotent group-scheme G is canonically identified to the completion of the lie sub-algebra of $k\{\{X\}\}$ generated by $X_1, X_2, \cdots, X_g$. This is an inverse limit of finite dimensional nilpotent lie algebras and the inverse limit of the corresponding nilpotent algebraic groups is precisely G.

In positive characteristic, the situation is more difficult. Even in the commutative case, a complete classification of such objects is given by Deindorne modules. We do not know as yet the non-commutative analogue of this. Essentially it amounts to a classification of finite nilpotent group-schemes.

Here are some elementary examples :

Example 1 : Take disjoint sets $S_1, S_2, \cdots, S_t$ of k-rational points of $\mathbf{P}'$ such that the sum of the cardinalities of the S_i is $g+t$. Let X be the "semi-normal" curve obtained by identifying all points of S_i to a single point y_i for $i = 1, 2, \cdots, t$.

Then $H^1(X, \mathbf{O}_X)$ has rank g and the associated formal group is given by $F_i(X;Y) = X_i + Y_i + X_i Y_i$.

X is semi-normal if the local ring completions of X are isomorphic to $k[[x_1, x_2, \cdots, x_n]]/(x_i x_j$ for all $i < j)$, for some n.

Proof : For simplicity assume that $S = S_1$ and $t = 1$, and $x_0, x_1, \cdots, x_g$ are the points of S. The image of x_0 in X will still be denoted by x_0.

If $f: \mathbf{P}^1 \to X$ is the normalisation map and W is in $N(X)$, then $f^*(W)$ is in $N(\mathbf{P}^1)$ and is therefore trivial. Thus $f^*W = V \otimes_k \mathbf{O}_{\mathbf{P}^1}$ canonically, where V is the fibre of $f^*(W)$ at x_0. The diagram

$$\begin{tikzcd}[column sep=small] & \mathbf{P}^1 \arrow[dr, "f"] & \\ \operatorname{Spec} k \arrow[ur, "x_i"] \arrow[dr, "x_0"'] & & X \\ & \mathbf{P}^1 \arrow[ur, "f"'] & \end{tikzcd}$$

gives an isomorphism of the fibres of $f^*(W)$ at x_0 and x_i for $1 \leqslant i \leqslant g$, *i.e.*, an automorphism ϕ_i of V for $1 \leqslant i \leqslant g$.

Using the filtration that W possesses, we see that there is a flag $V = V_0 \supseteq V_1 \supseteq V_2 \cdots \supseteq V_m = 0$ so that $\phi_i X - X \in V_{j+1}$ for all $x \in V_j$. Consequently $\phi_j - I_v = \psi_i$ has the property that $\psi_{i_0} \psi_{i_1} \cdots \psi_{i_m} = 0$ for all possible choices of $i_0, i_1, \cdots, i_m \in \{1, 2, \cdots, g\}$. This makes V a $k\{\{X\}\}$-module by letting the X_i act by ψ_i. This immediately gives :

$$\begin{tikzcd} k\{\{X\}\}\text{-mod} \arrow[rr, leftrightarrow, "\cong"] \arrow[dr] & & N(X) \arrow[dl, "T"] \\ & k\text{-mod} & \end{tikzcd}$$

where the vertical arrows are the forgetful functor and evaluation at x_0 respectively. We have a natural isomorphism of A with $k\{\{X\}\}$ in this case (unlike the necessarily arbitrary isomorphism of Proposition 4).

If W' and W'' are in $N(X)$ with the corresponding ϕ'_i and ϕ''_i, then $W' \otimes W''$ is defined by $\phi'_i \otimes \phi''_i$. Putting $\phi'_i = 1 + \psi'_i$ and $\phi''_i = 1 + \psi''_i$, this gives $\phi_i' \otimes \phi''_i = 1 + \psi'_i \otimes 1 + 1 \otimes \psi''_i + \psi'_i \times \psi''_i$, and therefore the homomorphism $k\{\{X\}\} \to k\{\{X\}\} \hat{\otimes} k\{\{X\}\}$ is given by $X_i \to X_i \otimes 1 + 1 \otimes X_i + X_i \otimes X_i$.

In our language, $F_i(X ; Y) = X_i + Y_i + X_i Y_i$. Q.E.D.

Example 2 : If $y_1, y_2, \cdots, y_g$ are k-rational points of $\mathbf{P}'$ and X is obtained from $\mathbf{P}'$ by introducing a simple cusp at $y_1, y_2, \cdots, y_g$ then $rk\, H^1(X, \mathbf{O}_X) = g$, and $F(X ; Y) = X_i + Y_i$ in this case.

This means that $f : \mathbf{P}^1 \to X$ is set—theoretically injective and if $f(y_i) = x_i$, then $\mathbf{O}_{X, x_i} = k + m_i^2$ where m_i is the maximal ideal of $\mathbf{O}_{\mathbf{P}^1, y_i}$.

Proof : Choose a point $x_0 \in \mathbf{P}^1$. Let W be in $N(X)$. As before $f^*W = V \otimes_{\mathbf{P}}$ where V is the fibre of f^*W at x_0. Thus all fibres can be canonically identified to V. The diagram :

$$\begin{tikzcd} \operatorname{Spec} k[q/q^2] \arrow[r] \arrow[d] & \mathbf{P}' \arrow[d, "f"] \\ \operatorname{Spec} k \arrow[r, "x_i"] & X \end{tikzcd}$$

gives an automorphism of $V \otimes k[q/q^2]$ which is the identity modulo q. Thus this automorphism ϕ_i is of the form $1 + q\psi_i$ where $\psi_i \in \operatorname{End}(V)$. Using the filtration of W, the ψ_i preserve a flag as in example 1, showing that V becomes a $k\{\{X\}\}$-module.

If W' and W'' are in $N(X)$ with the ϕ'_i and ϕ''_i, then $\phi'_i \otimes \phi''_i = 1 + q(\psi'_i \otimes 1 + 1 \otimes \psi''_i)$ and this shows that

$$F(X\,;\,Y) = X_i + Y_i.$$

Example 3 : What is the affine group-scheme G associated to $F(X\,;\,Y) = X_i + Y_i + X_i\, Y_i$? In characteristic zero, we have already seen the answer, so we restrict ourselves to characteristic p with $p > 0$. Here the answer is : G is isomorphic to a free pro-p-group on g letters. In particular, G is reduced.

Proof : Let π be a free group on $u_1, u_2, \cdots, u_g$ and let P be its pro-p-completion.

Let $A(H)$ be the group algebra of a finite group H. Then $A(P) =$ dual of its co-ordinate ring $= \varprojlim_H A(H)$ where H runs through all finite p-quotients H of π.

Let $\rho : k\{X_1, X_2, \cdots, X_g\} \to A(P)$ be the homomorphism defined by $\rho(X_i) = u_i - 1$. For any finite p-quotient H of π, the composite $k\{X_1, X_2, \cdots, X_g\} \to A(P) \to A(H)$ is clearly surjective, andits kernel contains $(X_1, X_2, \cdots, X_g)^h$ where h is the cardinality of H. This is seen as usal by showing that $A(H)$ has a flag such that u_i acts by the identity on successive quotients. Therefore there is an induced diagram :

$$\begin{tikzcd}
k\{X_1, X_2, \ldots, X_g\} \arrow[r, "\rho"] \arrow[d] & A(P) \\
k\{\{X\}\} = k\{\{X_1, \ldots, X_g\}\} \arrow[ur, "\bar{\rho}"'] &
\end{tikzcd}$$

and $\bar{\rho}$ is a surjection.

Now consider $B_n = |\, k\{\{X\}\}\,|$ the n-th power of its maximal ideal. The image of the homomorphism $\pi \mapsto B_n^*$ given by $u_i \mapsto 1 + X_i$ is denoted by H_n. We shall show below (lemma 9) that H_n is a finite p-group. This induces $A(P) \to A(H_n) \to B_n$ for all n, and therefore a continuous homomorphism $A(P) \to \varprojlim_n B_n = k\{\{X\}\}$. This is seen to be the inverse of $\bar{\rho}$ quite easily.

The diagonal homomorphism $A(P) \to A(P) \hat{\otimes} A(P)$ is given by $u_i \to u_i \otimes u_i$. Under $\rho : k\{\{X\}\} \xrightarrow{\cong} A(P)$, this becomes $X_i \to X_i \otimes 1 \oplus 1 \otimes X_i \oplus X_i \otimes X_i$. This finishes the proof modulo the following well-known.

Lemma 9 : Any finitely generated subgroup of a nilpotent affine algebraic group in characteristic p is a finite p-group.

Proof : By induction on the dimension of the algebraic group N : there is an exact sequence $1 \to N_1 \to N \to G_a \to 1$. If H is the finitely generated subgroup, then $1 \to N_1 \cap H \to H \to G_a$ is exact. The image of H in G_a is a finite abelian p-group. Consequently $N_1 \cap H$ is of finite index in H and is therefore finitely generated. By induction, $N_1 \cap H$ is a finite p-group and therefore H itself is a finite p-group.

We need the following :

Lemma 10 : The letters A and B stand for inverse limits of finite dimensional local k-algebras with residue field k. It will also be assumed that they have finite

dimensional m/m^2. Any such algebra will be called free if it is isomorphic to $k\{\{X_1, X_2, \cdots, X_g\}\}$ for some g.

A. If $f : A \to B$ is a homomorphism inducing a surjection on the m/m^2-level, then f is a surjection.

B. A homomorphism $f : A \to B$ inducing an isomorphic on the m/m^2-level is an isomorphism if there is a $g : B \to A$ such that $fg\,(X) = X$ for all $X \in B$.

C. If $f : A \to B$ induces an isomorphism on the m/m^2-level and B is free, then f is an isomorphism.

D. If A is free and $h_1, h_2, \cdots, h_r \in m$ are such that their images in m/m^2 are inearly independent, then $A/(h_1, h_2, \cdots, h_r)$ is free.

Proof; A. The hypothesis implies that there is a surjection on the m^n/m^{n+1} level for all n and this is enough.

B. Clearly g is injective. But g induces an isomorphism on the m/m^2-level implying by A that it is also surjective.

C. By Part A of the lemma, $f : A \to B$ is surjective. But B is free ; therefore there is a $g : B \to A$ such that $fg\,(x) = x$ for all $x \in B$. By Part B of the lemma, f is an isomorphism.

D. Choose $g_1, g_2, \cdots g_s$ in the maximal ideal so that the h_i anl the g_j form a basis for m/m^2. Let B be a free algebra on $r + s$ generators and define $f : B \to A$ by sending the generators to the h_i and the g_j. By Part C, f is an isomorphism, therefore

$$A/(h_1, h_2, \cdots, h_r) \cong k\{\{X_1, X_2, \cdots, X_{r+s}\}\}/(X_1, X_2, \cdots, X_r) \cong k\{\{Y_1, Y_2, \cdots, Y_s\}\}.$$

This proves the lemma completely.

Lemma 11 : Let G be the affine group-scheme associated to a non-commutative formal group. Assume that k is perfect. Then G_{red} is also a group-scheme associated to a non-commutative formal group. If k is algebraically close then, G_{red} is a free pro-p-group.

Proof: Let A and B be the duals of the co-ordinate rings of G and G_{red} respectively. By assumption A is free, and the first assertion of the lemma is equivalent to the assertion that B is free.

The exact sequence $1 \to G_{loc} \to G \to G_{red} \to 1$ is split by the natural inclusion of G_{red} in G. This gives $j : A \to B$ and $i : B \to A$ so that $j\,i\,(x) = x$ for all $x \in B$. Choose $h_1, h_2, \cdots, h_r$ in the kernel of j so that their images in ker $(j) + m^2/m^2$ form a basis. Let $p : A \to A/(h_1, h_2, \cdots, h_r) = C$ be the projection. If $f \circ p = j$ and $p \circ i = g$, then $fg\,(x) = x$ for all $x \in B$. Also C is free by Lemma 10·D. By 10·B, f is an isomorphism, showing that B is free.

If k is algebraically closed, there is a surjection $h : F \to G_{red}$ where F is a free pro-p-group such that Hom $(G_{red}, Z/p) \to$ Hom $(F, Z/p)$ is an isomorphism. Consequently, $A(F) \to B$ induces an isomorphism on the m/m^2-level. But B is free and by 10·C it follows that $A\,(F) \to B$ is an isomorphism. Therefore $F \to G_{red}$ is an isomorphism.

Corollary (due to Safarevich) : Let X be a complete curve with $\Gamma(X, O_x) = k$. The maximal p-quotient of the etale fundamental group is a free pro-p-group in characteristic p.

If $f: Y \to X$ is a Galois etale covering of degree N where N is a power of p, and $r(Y)$ and $r(X)$ are the ranks of the Hasse-Witt matrices of Y and X respectively, then

$$r(Y) - 1 = N(r(X) - 1).$$

We first remark that Safarevich's proof is much shorter than ours.

Proof : In this set-up the maximal p-quotient of the etale fundamental group is just $U(X, x_0)_{\mathrm{red}}$. By Lemma 11 and Proposition 4, this is a free pro-p-group. The second assertion is an immediate consequence.

We state without proof the following proposition which shows that $U(X, x_0)$ has flat variation at least in a special case :

Proposition 7 : Let $f: X \to S$ be a flat proper morphism with fibres of dimension one with $S = \operatorname{Spec} R$. Let $j: S \to X$ be a section. There is an affine group-scheme $U(X, j)$ which is S-flat such that for all $t : \operatorname{Spec} k \to S$, $U(X, j)_t = U(X_t, x_0)$ where x_0 is the base-point of X_t induced by j.

We assume that $f_* \mathbf{O}_s = \mathbf{O}_s$ and $R^1 f_* \mathbf{O}_s$ is locally free. If we assume further that $R^1 f_* \mathbf{O}_s = \mathbf{O}_s^g$, then this gives a non-commutative formal group $F(X:Y)$ with coefficients in R.

The proof is not difficult ; one has to construct the V_n as in Lemma 7 for this situation and construct the A/J_n as $\operatorname{End}(V_n)^\circ$.

We proved that $U(X, x_0) \times U(Y, y_0) = U(X \times Y, x_0 \times y_0)$. We conjecture that the same holds for the fundamental group-scheme. This would show that $\pi(X, 0)$ is abelian for an abelian variety X and this shows in fact that $\pi(X, 0) = \varprojlim_{G \subset \to \hat{X}} \hat{G}$ where G ranges through all finite subgroup-schemes of $\hat{X}$. The best we can manage now is :

Proposition 8 : If k is perfect and X is an elliptic curve, then $\pi(X, 0) = \varprojlim_{G \to \hat{X}} \hat{G}$.

Proof : By Proposition 5, Chapter II, k may be replaced by any separable extension and therefore we may assume it is algebraically closed.

With the V_n as in Lemma 7, a theorem of Atiyah asserts that any semi-stable bundle of degree zero is a direct sum of bundles of the type $L \otimes V_n$ where L is a line bundle on X.

This gives us an easy classification of all essentially finite bundles on X : direct sums of $L \otimes V_n$ with L ranging through all line bundles of finite order.

There is a natural surjection $\pi(X, 0) \to \varprojlim_{G \to \hat{X}} \hat{G}$ and the representations of the group on the right already give all essentially finite bundles : if L is a line bundle of order m, this gives a $Z/m \to \hat{X}$, and the V_n come from representations of $U(X, 0)$ which by the Corollary to Lemma 8 is a quotient of $\varprojlim_{G \to \hat{X}} \hat{G}$.

Therefore any representation of $\pi(X, 0)$ is already a representation of $\varprojlim_{G \to \hat{X}} \hat{G}$ showing that the homomorphism is indeed an isomorphism.

We examine finally the behaviour of $U(X, x_0)$ under base-change. Let L be any field extension of k and $(\bar{X}, \bar{x}_0)$ and $\overline{U(X, x_0)}$ be the base-change of (X, x_0) and $U(X, x_0)$ to L respectively. The universal properties show that there is a natural homomorphism $U(\bar{X}, \bar{x}_0) \to \overline{U(X, x_0)}$ with a commutative diagram :

$$\begin{tikzcd} U(X, x_0)-\text{mod} \arrow[r] \arrow[d, "="', "F"] & U(\bar{X}, \bar{x}_0)-\text{mod} \arrow[d, "="', "\bar{F}"] \\ N(X) \arrow[r] & N(\bar{X}) \end{tikzcd}$$

where the horizontal arrows are the obvious ones and F is the given natural equivalence of categories.

Let B be the dual of the co-ordinate ring of $U(X, x_0)$ and let I_n be the closure of the n-th power of its maximal ideal. The arrow $U(\bar{X}, \bar{x}_0) \to \overline{U(X, x_0)}$ induces a continuous homomorphism $f : B \to A = A \otimes_k L$, and the first horizontal arrow in the above commutative diagram is induced by this homomorphism.

If W_n is defined inductively by $W_{n+1} = U(W_n)$ and $W_1 = \mathbf{O}_{\bar{x}}$, then $W_n = \bar{F}(B/I_n)$. But, by induction, it follows that $\bar{V}_n$ (which is the base-change of V_n to $\bar{X}$) is isomorphic to W_n, because $H^1(X, \bar{V}_n^*) = H^1(X, V_n^*) \otimes_k L$.

By the above commutative diagram, it follows that $(A/J_n) \otimes_k L$ considered as a B-module is isomorphic B/I_n, i.e., $f^{-1}(J_n \otimes L) = I_n$ and $B/I_n \to A \otimes_k L/J_n \otimes_k L$ is subjective for all n. From this, $f : B \to A$ is itself an isomorphism, showing that

Proposition 9 : With notation as above, $U(\bar{X}, \bar{x}_0) \to \overline{U(X, x_0)}$ is an isomorphism. In other words, $U(X, x_0)$ is invariant under base-change.

Appendix

Tannaka categories

Section 2 of this appendix contains the proofs of all the results about Tannaka categories that have been used freely (see chapter I, § 1). These follow easily enough from the results of § 1.

Section 1. Let k be a field. The only algebras considered here are k-algebras A equipped with a topology such that

(a) A/J is a finite dimensional k-vector space for all open two-sided ideals J, and

(b) $A \to \varprojlim_{J} A/J$ is an isomorphism where the J run through all open two-sided ideals in A.

In particular any finite dimensional k-algebra with the discrete topology will do

All k-algebra homomorphisms under consideration will be assumed to be continuous.

If A and B are such algebras, $A \otimes B$ has a topology by taking $\{ I \otimes B + A \otimes J :$ I and J are open two-sided ideals of A and B respectively$\}$ to be a basis of

The category of modules over a ring is denoted both by A–mod and | A | in the text.

neighbourhoods of zero. The completion $\varprojlim_{I,J} A \otimes B/I \otimes B + A \otimes J = \varprojlim_{I,J} A/I \otimes B/J$ will be denoted by $A \hat{\otimes} B$. Here a neighbourhood-basis consists of the kernels of $A \hat{\otimes} B \to A/I \otimes B/J$ with I and J as usual.

By $| A |$ we shall mean the category of left A-modules M which are finite-dimensional k-vector spaces and whose annihilators are open.

A homomorphism $f : A \to B$ induces a functor $H(f) : | B | \to | A |$. If $i_A : k \to A$ is the canonical inclusion, we put $H(i_A) = T_K$.

More generally, if $f : A \to B_1 \hat{\otimes} B_2 \hat{\otimes} \cdots \hat{\otimes} B_n$ is a homomorphism, then there is a functor $H(f) : | B_1 | \times | B_2 | \times \cdots \times | B_n |$.

The main result of this section is :

Proposition 1 : Let $\mathcal{C}$ be an abelian category with finite direct sums. Assume that $\mathcal{C}(V, W) = \mathcal{C}$-morphisms from V to W, where V and W are objects of $\mathcal{C}$ has the structure of a k-vector space and that for objects V, W, P of $\mathcal{C}$, the composition $\mathcal{C}(V, W) \times \mathcal{C}(W, P) \to \mathcal{C}(V, P)$ is k-bilinear, Such a $\mathcal{C}$ will be called an abelian k-category.

Assume further that $\operatorname{Obj} \mathcal{C}$ is a set, and that $T : \mathcal{C} \to | k |$ is a faithful exact k-linear functor. The phrase "k-linear" in this context means that $T(V, W) : \mathcal{C}(V, W) \to \operatorname{Hom}_k(TV, TW)$ is k-linear for all objects V and W of $\mathcal{C}$.

There is then an algebra $A(\mathcal{C})$ and an equivalence $F : \mathcal{C} \to | A(\mathrm{C}) |$ with the commutative diagram :

$$\begin{tikzcd} \mathcal{C} \arrow[rr, "F"] \arrow[dr, "T"'] & & A(\mathcal{C})\text{-mod} \arrow[dl, "T_k"] \\ & k\text{-mod} & \end{tikzcd}$$

The proof of this Proposition will take up the rest of this section.

The construction of $A(\mathcal{C})$:

This is forced on us. Suppose we are given a commutative diagram :

$$\begin{tikzcd} \mathcal{C} \arrow[rr, "R"] \arrow[dr, "T"'] & & B\text{-mod}, \arrow[dl, "T_k"] \\ & k\text{-mod} & \end{tikzcd}$$

Then, for all objects V of $\mathcal{C}$, TV becomes a B-module in a natural way and therefore there is a homomorphism $\rho_V : B \to \operatorname{End}(TV)$ vanishing on some open two-sided ideal.

If $f \in \mathcal{C}(V, W)$, then $Tf : TV \to TW$ is B-linear showing $Tf \circ \rho_V(b) = \rho_W(b) \circ Tf$ for all $b \in B$.

Put $\rho = \prod_V \rho_V : B \to \prod_{V \in \operatorname{Obj} \mathrm{C}} \operatorname{End}(TV)$ and let $\pi_W : \prod_{V \in \operatorname{Obj} \mathrm{C}} \operatorname{End}(TV) \to \operatorname{End}(TW)$ be the projection for each $W \in \operatorname{Obj} \mathcal{C}$.

Then $Tf \circ \pi_V(\rho(b)) = \pi_W(\rho(b)) \circ Tf$.

Therefore, if $A(\mathcal{C}) = \{a \in \prod_{V \in \operatorname{Obj} \mathrm{C}} \operatorname{End} TV \mid \forall f \in \mathcal{C}(V, W), \forall V \in \operatorname{Obj} \mathcal{C}, \forall W \in \operatorname{Obj} \mathcal{C}, Tf \circ \pi_V(a) = \pi_W(a) \circ Tf\}$, the image of ρ is contained in $A(\mathcal{C})$,

Furthermore, if $\prod_{V \in \mathrm{Obj}\, C} \mathrm{End}\ TV$ is given the product topology (with all finite dimensional spaces having the discrete topology), clearly $B \to A(\mathcal{C})$ is continuous.

Finally, for all $V \in \mathrm{Obj}\, \mathcal{C}$, $A(\mathcal{C}) \to \prod_{V \in \mathrm{Obj}\, C} \mathrm{End}\ TV \xrightarrow{\pi_V} \mathrm{End}\ TV$ makes TV an object of $|A(\mathcal{C})|$ which we shall denote by FV. All this goes to show that we have a commutative diagram :

$$\begin{tikzcd} \mathcal{C} \arrow[rr, "F"] \arrow[dr, "T"'] & & A(\mathcal{C})\text{-mod} \arrow[dl, "T_k"] \\ & k\text{-mod} & \end{tikzcd}$$

and that any other commutative diagram

$$\begin{tikzcd} \mathcal{C} \arrow[rr, "R"] \arrow[dr, "T"'] & & B\text{-mod} \arrow[dl, "T_k"] \\ & k\text{-mod} & \end{tikzcd}$$

is induced by a unique $\rho : B \to A(\mathcal{C})$; in other words, $R = H(\rho) \circ F$.

If $(\mathcal{C}, T) = (|A|, T_k)$ it is easy to see that $A = A(\mathcal{C})$ canonically. This proves:

Corollary : a $F : |A| \to |B|$ such that $T_k \circ F = T_k$ is equal to $H(\rho)$ for a unique $\rho : B \to A$. However we need the following slightly stronger statement for §2.

Proposition 2 : Let $\otimes^n : |k| \times |k| \times \cdots \times |k| \to |k|$ be the usual tensoring functor. The functors $F : |B_1| \times \times |B_2| \times \cdots \times |B_n| \to |A|$ such that $T_k \circ F = \otimes^n \circ (T_k \times T_k \times \cdots \times T_k)$ are in one-one correspondence with $f : A \to B_1 \hat{\otimes} B_2 \hat{\otimes} \cdots \hat{\otimes} B_n$.

Proof : For ease of writing, we take $n = 2$.

We may ignore T_k if we agree to identify modules with their underlying vector spaces. What F does is the following :

A. For objects M_1 and M_2 of $|B_1|$ and $|B_2|$, there is an A-module structure on $M_1 \otimes M_2$. For $a \in A$ and $m \in M_1 \otimes M_2$, the multiple of m by a will be denoted by $a \cdot m$.

B. If $f_1 : M_1 \to N_1$ and $f_2 : M_2 \to N_2$ are module homomorphisms for B_1 and B_2 respectively, then $f_1 \otimes f_2$ is a A-module homomorphism.

In particular, putting $B_1/J_1 = M_1 = N_1$ and $B_2/J_2 = M_2 = N_2$ for two-sided open ideals J_1 and J_2, and f_i = right multiplication by $b_i \in B_i/J_i$ for $i = 1$ and 2, we see that the A-action on $B_1/J_1 \otimes B_2J_2$ commutes with all right multiplications. Consequently there is a homomorphism $h(J_1, J_2) : A \to B_1/J_1 \otimes B_2/J_2$ such that $a \cdot m = (h(J_1, J_2)\, a)m$ for all $a \in A$, $m \in B_1/J_1 \otimes B_2/J_2$.

It is easy enough to see that the $h(J_1, J_2)$ form an inverse system giving rise to $h : A \to B_1 \hat{\otimes} B_2$, such that the composite

$$\begin{tikzcd} A \arrow[r, "h"] \arrow[rr, bend right, "h(J_1, J_2)"'] & B_1 \hat{\otimes} B_2 \arrow[r] & B_1/J_1 \otimes B_2/J_2 \end{tikzcd}$$

We shall show that $F = H(h)$, *i.e.*, for all $m \in M_1 \otimes M_2$, $a \in A$, $a \cdot m = h(a)\, m$. If M_1 and M_2 are annhilated by J_1 and J_2 respectively, define $f_i : B_i/J_i \to M$ by $f_i(b) = bm_i$ for $i = 1, 2$, Then $a \cdot (m_1 \otimes m_2) = a \cdot ((f_1 \otimes f_2(1 \otimes 1)) = (f_1 \otimes f_2)(a \cdot (1 \otimes 1)) = (f_1 \otimes f_2)(h(a)) = h(a)(m_1 \otimes m_2)$.

But the $m_1 \otimes m_2$ generate all of M so that $a \cdot m = h(a)\, m$ always. Therefore $F = H(h)$.

Given $f : A \to B_1 \otimes B_2$ and $F = H(f)$ it is easy to see that the h constructed above equals f. This establishes the one-to-one correspondence and completes the proof of Proposition 2.

We have already constructed $F : \mathcal{C} \to | A(\mathcal{C}) |$. To show that this is an equivalence, we must prove :

*F*1. For all V and W in $\mathcal{C}$, $\mathcal{C}(V, W) \to \mathrm{Hom}_{A(\mathcal{C})}(FV, FW)$ is an isomorphism. Note that it is already a monomorphism because $T = T_k \circ F$ is faithful.

*F*2. For every M in $| A(\mathcal{C}) |$ there is a V in $\mathcal{C}$ such that $FV = M$.

What we need is functors going the other way :

Lemma 1 : Let B be finite dimensional k-algebra. Diagrams

$$\begin{tikzcd} B\text{-mod} \arrow[rr, "G"] \arrow[dr] & & \mathcal{C} \arrow[dl] \\ & k\text{-mod} & \end{tikzcd}$$

are in one-one correspondence with the data :

1. An object N of $\mathcal{C}$ and an isomorphism $TN \cong B$,
2. a k-linear $\rho : B \to \mathcal{C}(N, N)$ such that under the above isomorphism $T_\rho(b)$ = right multiplication by $b \in B$.

Proof : Given G, put $GB = N$. Then $B = T_k GB = TN$, and if $R_b : B \to B$ is right multiplication by $b \in B$, put $\rho(b) = G(R_b)$.

Conversely given N, the isomorphism $TN \cong B$, and $\rho : B \to \mathcal{C}(N, N)$. First note that $\rho(a) \circ \rho(b) = \rho(ba)$ for all $a, b \in B$. This is so because T is faithful and $T(\rho(a)\rho(b)) = R_a R_b = R_{ba} = T(\rho(ba))$.

Let P be any object of $| B |$. Fix a presentation :

$$B^p \xrightarrow{h} B^q \to P \to 0.$$

If h is given by the matrix (h_{ij}), define GP by an exact sequence in $\mathcal{C}$:

$$N^p \xrightarrow{\rho(h)} N^q \to GP \to 0$$

where $\rho(h)$ is the matrix with entries $\rho(h_{ij})$.

If $f : P \to Q$ is a B-module homomorphism and $B^s \xrightarrow{h'} B^q \to Q \to 0$ is the chosen presentation of Q, then there is a diagram :

$$\begin{tikzcd} B^p \arrow[r, "h"] \arrow[d, "r'"] & B^q \arrow[r] \arrow[d, "r"] & P \arrow[r] \arrow[d, "f"] & 0 \\ B^r \arrow[r] & B^s \arrow[r] & Q \arrow[r] & 0 \end{tikzcd}$$

which in turn induces a unique diagram :

$$\begin{tikzcd} N^p \arrow[r, "\rho(h)"] \arrow[d, "\rho(r')"] & N^q \arrow[r] \arrow[d, "\rho(r)"] & GP \arrow[r] \arrow[d, "l"] & 0 \\ N^r \arrow[r, "\rho(h')"] & N^s \arrow[r] & GQ \arrow[r] & 0 \end{tikzcd}$$

Now $l : GP \to GQ$ does not depend on the choice of r and r' because $Tl : TGP \to TGQ$ is just $T_k f : T_k P \to T_k Q$ and T is faithful. Put $Gf = l$.

This defines G. We omit to check that G is a k-linear functor and the fact that this establishes a one-to one correspondence with the G and the N with the above data.

Lemma 2 : With N and G as above, there is a unique $f : A(\mathcal{C}) \to B$ such that $F \circ G = H(f)$. Thus if P and Q are B-modules, $V = GP$ and $W = GQ$, then $FV = H(f)P$ and $FW = H(f)\,Q$ and the image of $F : \mathcal{C}(V, W) \to \mathrm{Hom}_{A(\mathcal{C})}(FV, FW)$ contains the image of $\mathrm{Hom}_B(P, Q) \to \mathrm{Hom}_{A(\mathcal{C})}(FV, FW)$.

Proof : The existence and uniqueness of f follows from Proposition 2. The next assertion follows from the diagram :

$$\begin{tikzcd} \mathrm{Hom}_B(P, Q) \arrow[r, "G(P,Q)"] \arrow[rr, bend right, "H(f)(P,Q)"'] & \mathcal{C}(V, W) \arrow[r, "F(V,W)"] & \mathrm{Hom}_{A(\mathcal{C})}(FV, FW) \end{tikzcd}$$

We shall use this lemma while proving $F1$.

Lemma 3 : With N and G and B as in Lemma 1, we shall characterise GP for P in $|B|$.

Given (a) $V \in \mathrm{Obj}\,\mathcal{C}$

(b) $h : P \to \mathcal{C}(N, V)$

a k-linear map such that $h(ap) = h(p) \circ \rho(a)$ for all $a \in B$, so that

(c) $\tilde{h} : P \to TV$ defined by $\tilde{h}(p) =$ the value at 1 of $T(h(p)) : TN \cong B \to TV$ for all $p \in P$, is an isomorphism, then $V = GP$. In future, GP will be denoted by $N \otimes_B P$.

This is fairly obvious so we skip the proof.

Lemma 4 : Let N and B be as in Lemma 1. Let $\phi : B \to C$ be an algebra homomorphism with C finite dimensional. Let Q be a left C-module. Then

(a) there is an anti-homomorphism $\tilde{\rho} : C \to \mathcal{C}(N \otimes_B C, N \otimes_B C)$ and an isomorphism $\mathcal{C} \cong T(N \otimes_B C)$ so that $T(c)$ is right multiplication by c for all $c \in C$ under this isomorphism.

(b) $(N \otimes_B C) \otimes_C Q \cong N \otimes_B Q$ with Q considered as a B-module in an obvious manner.

Proof (a) : With G as in Lemma 1, the anti-homomorphism $C \to \mathrm{End}_C(C, C) \to \mathrm{End}_B(C, C)$ induces the anti-homomorphism $\tilde{\rho} : C \to \mathcal{C}(GC, GC)$. The rest of (a) follows from the fact that $T_k = T \circ G$.

(b) Let $V = (N \otimes_B C) \otimes_C Q$. Then there is a k-linear $h : Q \to (N \otimes_B C, V)$ such that $h(cq) = h(q) \circ \tilde{\rho}(c)$ for all $c \in C$, $q \in Q$. Using $G\varphi : GB = N \to GC = N \otimes_B C$, we define $h'(q) = h(q) \circ G\phi$. This gives $h' : Q \to \mathcal{C}(N, V)$. We check that h' has the desired properties :

1. $$\begin{aligned} h'(q) \circ \rho(b) &= h(q) \circ G\phi \circ G(R_b) \\ &= h(q) \circ \tilde{\rho}(\varphi b) \circ G\phi \\ &= h(\phi(b)q) \circ G\phi \\ &= h'(\phi(b)q). \end{aligned}$$

2. We need to show that $q \mapsto (T h'(q))(1)$ is an isomorphism from Q to $T((N \otimes_B C) \otimes Q)$. But this is the same as $q \mapsto (T h(q))(1)$ because there is a commutative diagram

$$\begin{tikzcd} TGB \arrow[r, "T\phi"] \arrow[d] & TGC \arrow[d] \\ B \arrow[r, "\phi"] & C \end{tikzcd}$$

which takes 1 to 1. This finishes the proof of the lemma.

Definition : A pair (B, N) with the data as in Lemma 1 will be called a ring object in $\mathcal{C}$. For a P in $|B|$, GP will be denoted by $N \otimes_B P$.

We shall now construct plenty of ring objects in $\mathcal{C}$. For any subset S of Obj $\mathcal{C}$, let $C(S) = \prod_{V \in S} \mathrm{End}(TV)$ and let $\pi_V : C(S) \to \mathrm{End}(TV)$ be the projection for $V \in S$.

Let $A(S) = \{a \in C(S) \mid \forall\, V \in S, \forall\, W \in S, \forall f \in \mathcal{C}(V, W), Tf \circ \pi_V(a) = \pi_W(a) \circ Tf\}$

There is a natural homomorphism $A(\mathcal{C}) \to A(S)$.

All subsets of Obj $\mathcal{C}$ considered from now on will be assumed to be finite.

Lemma 5 : $A(\mathcal{C}) \to \varprojlim_S A(S)$ is an isomorphism.

This is obvious.

Lemma 6 : For each finite $S \in \text{Obj } \mathcal{C}$, let $\tilde{A}(S)$ be the image of $A(\mathcal{C})$ in $A(S)$. Then

(a) there is a finite T containing S such that $A(T) \to A(S)$ has its image equal to $\tilde{A}(S)$, and

(b) $A(\mathcal{C}) = \varprojlim_{S} \tilde{A}(S)$.

Proof : (b) is clear so we need to prove only (a).

Let $A_T(S)$ be the image of $A(T) \to A(S)$ for all T containing S. There is some T containing S for which $\dim A_T(S)$ is the least possible. Consequently for all $T' \supseteq T$, $A_{T'}(S) \to A_T(S)$ is an isomorphism. Put $A_T(S) = X(S)$.

If $S_1 \subseteq S_2$, choose T_1 and T_2 containing S_1 and S_2 respectively with the above property. If $T = T_1 \cup T_2$, the diagram

$$\begin{tikzcd}[column sep=small] & A(T_1) \arrow[rrr] & & & A(S_1) \\ A(T) \arrow[ur] \arrow[dr] & & & & \\ & A(T_2) \arrow[rr] & & A(S_2) \arrow[uur] & \end{tikzcd}$$

shows that the image of $X(S_2)$ in $A(S_2) \to A(S_1)$ is precisely $X(S_1)$. Clearly, $\varprojlim_{S} X(S) \to \varprojlim_{S} A(S)$ is an isomorphism. But $\{X(S)\}$ is an inverse system of surjections showing that the image of $A(\mathcal{C}) \to A(S)$ is precisely $X(S)$. Therefore $X(S) = \tilde{A}(S)$ and this proves the lemma.

For every (finite) subset S of Obj $\mathcal{C}$ we shall construct ring objects $(A(S), B(S)$ and $(C(S), D(S))$.

We need first to make some trivial remarks: there is a unique functor Hom $|k| \times \mathcal{C} \to \mathcal{C}$ which is k-linear in each variable, contravariant in the first and covariant in the second variable, and which satisfies : $\mathcal{C} \to \mathcal{C}$ defined by $W \to \text{Hom}(k, W)$ is the identity functor.

Moreover there is a commutative diagram :

$$\begin{tikzcd} k\text{-mod} \times \mathcal{C} \arrow[r, "\text{Hom}"] \arrow[d, "1 \times T"'] & \mathcal{C} \arrow[d, "T"] \\ k\text{-mod} \times k\text{-mod} \arrow[r, "\text{Hom}"] & k\text{-mod} \end{tikzcd}$$

where the Hom in the second row is the usual one.

The category $|k|$ can be assumed to have objects k^n for $n = 0, 1, 2, \cdots$. Define $\text{Hom}(k^n, W) = W^n$ for all $W \in \text{Obj } \mathcal{C}$.

Given $f : k^n \to k^m$ and $g : W_1 \to W_2$ with W_1 and W_2 in $\mathcal{C}$, the corresponding homomorphism from $\text{Hom}(k^m, W_1) \to \text{Hom}(k^n, W_2)$, *i.e.*, from $W_1^m \to W_2^n$ is given by $\sum_{r,s} f_{rs}\, i_r \circ g \circ p_s$ where $i_r : W_2 \to W_2^n$ is the r-th inclusion,

$p_s : W_1^m \to W_1$ the s-th projection and the f_{rs} are the coefficients of the matrix f.

We have stated the Hom functor in basis-free language to avoid choosing bases for plenty of vector spaces which could be very cumbersome.

Definition : If $f : W_1 \to W_2$ is a linear transformation and $V \in \text{Obj } \mathcal{C}$, the $\mathcal{C}$ morphism from $\text{Hom}(W_2, V)$ to $\text{Hom}(W_1, V)$ will be denoted simply by R_f.

If $f \in \mathcal{C}(V_1, V_2)$ and W is a vector space, the $\mathcal{C}$-morphism from $\text{Hom}(W, V_1) \to \text{Hom}(W, V_2)$ will be denoted by L_f.

We now define the $D(S)$ and $B(S)$:

$D(S) = \bigoplus_{V \in S} \text{Hom}(TV, V)$. Clearly $TD(S) = C(S)$ canonically. For $V, W \in S$ and $f \in \mathcal{C}(V, W)$ we define $\alpha(f) \in \mathcal{C}(D(S), \text{Hom}(TV, W))$ by $\alpha(f) = L_f \circ p_V - R_{Tf} \circ p_W$ where the p_V and p_W are projections from $D(S)$ to $\text{Hom}(TV, V)$ and $\text{Hom}(TW, W)$ respectively.

Clearly $T(\alpha(f)) : C(S) \to \text{Hom}(TV, TW)$ is given by $T(\alpha(f))\, a = Tf \circ \pi_V(a) - \pi_W(a) \circ Tf$.

Define $B(S) = \bigcap_{f \in C(V, W);\, V, W \in S} \ker(\alpha(f))$. This makes sense because it is a finite intersection : any collection of f that span all the $\mathcal{C}(V, W)$ with V and W in S will do.

Clearly $TB(S) = A(S)$.

Lemma 7 : $(A(S), B(S))$ and $(C(S), D(S))$ are ring objects for all finite sets S. In addition,

1. $B(S) \otimes_{A(S)} C(S) = D(S)$
2. if $S_1 \subseteq S_2$, then $B(S_2) \otimes_{A(S_2)} A(S_1) = B(S_1)$
3. if $S_1 \subseteq S_2$, then $D(S_2) \otimes_{C(S_2)} C(S_1) = D(S_1)$.

Lemma 8 : If $S = \{V\}$ is a singleton, then TV is a $C(V) = \text{End}(TV)$-module for which $D(V) \otimes_{C(V)} TV = V$.

We first show that these lemmas together imply that F is an equivalance of functors. We first check $F2$.

Let M be an $A(\mathcal{C})$-module. Then M is an $\tilde{A}(S)$-module for some finite S and by lemma 6, $A(T) \to \tilde{A}(S)$ is a surjection for a suitable T containing S. Consequently M may be regarded as a $A(T)$-module. By lemma 2, $F(B(T) \otimes_{A(T)} M)$ is isomorphic to M. Therefore F induces a surjection from $\text{Obj } \mathcal{C}$ to $|\text{Obj}(A(\mathcal{C})|$,

Now we come to $F1$. For V and W in $\mathcal{C}$, we have to prove that $\mathcal{C}(V, W) \to \text{Hom}_{A(C)}(FV, FW)$ is a surjection. Let $S = \{V, W\} \subseteq \text{Obj } \mathcal{C}$. Then TV and TW are $A(S)$-modules in a natural manner and the corresponding $A(\mathcal{C})$-module structures induced by $A(\mathcal{C}) \to A(S)$ are precisely FV and FW respectively. Choose any $R \subset \text{Obj } \mathcal{C}$ which contains S. Let $S_1 = \{V\}$ and $S_2 = \{W\}$.

By Lemma 4 and Lemma 7, we have :

$$B(R) \otimes_{A(R)} TV = (B(R) \otimes_{A(R)} A(S_1)) \otimes_{A(S_1)} TV = B(S_1) \otimes_{A(S_1)} TV, \text{ and}$$

$$B(S_1) \otimes_{A(S_1)} TV = (B(S_1) \otimes_{A(S_1)} C(S_1)) \otimes_{C(S_1)} TV = D(S_1) \otimes_{C(S_1)} TV.$$

By Lemma 8, $D(S_1) \otimes_{C(S_1)} TV = V$. Thus we have : $B(R) \otimes_{A(R)} TV = V$ and $B(R) \otimes_{A(R)} TW = W$. By Lemma 2, the image of $\mathcal{C}(V, W) \to \text{Hom}_{A(C)}(FV, FW)$

contains all the $A(R)$-module homomorphisms from FV to FW. By choosing R large enough (by lemma 6), $A(R) \to A(S)$ has its image equal to $\tilde{A}(S)$, and for such R the $A(R)$-module homomorphisms and $A(\mathcal{C})$-module homomorphism are the same. This finishes the proof of Proposition 1 modulo Lemmas 7 and 8.

We retain the use of R_f and L_f defined by the Hom functor.

To define $\rho: C(S) \to \mathcal{C}(D(S), D(S))$, first consider the $\mathcal{C}$-morphism $R_{\pi_v(a)}$; $\mathrm{Hom}(TV, V) \to \mathrm{Hom}(TV, V)$ for all $V \in S$ and put $\rho(a) = \bigoplus_V R_{\pi_v(a)}$. With the natural identification of $TD(S)$ with $C(S)$ clearly $T\rho(a)$ is right multiplication by a.

This proves that $(C(S), D(S))$ is a ring object.

To show that $(A(S), B(S))$ is a ring object, it suffices to show that $\rho(a) B(S) \subseteq B(S)$ for all $a \in A(S)$.

Take any $a \in C(S)$. Then

$$\begin{aligned}
&\rho(a) B(S) \subseteq B(S) \\
&\Leftrightarrow 0 = \rho(a) B(S) + B(S)/B(S) \\
&\Leftrightarrow 0 = T(\rho(a) B(S) + B(S)/B(S)) \text{ by the faithfulness of } T \\
&\qquad = A(S)\, a + A(S)/A(S) \text{ by the exactness of } T \\
&\Leftrightarrow a \in A(S).
\end{aligned}$$

Therefore $(A(S), B(S))$ is a ring object.

The k-linear map : $C(S) \to \mathcal{C}(D(S), D(S)) \to \mathcal{C}(B(S), D(S))$ after an application of T becomes $D(S) \to \mathrm{Hom}_k(A(S), C(S))$ which is just $a \mapsto$ the restriction of the right multiplication by a to $A(S)$, for all $a \in C(S)$. This proves 7·1 :

$$B(S) \otimes_{A(S)} C(S) = D(S).$$

7·2 and 7·3 are equally clear : look at

$$C(S_2) \to \mathcal{C}(D(S_2), D(S_2)) \to \mathcal{C}(D(S_2), D(S_1)) \text{ and}$$
$$A(S_2) \to \mathcal{C}(B(S_2), B(S_2)) \to \mathcal{C}(B(S_2), B(S_1)) \text{ given by}$$

composing with the projections $D(S_2) \to D(S_1)$ and $B(S_2) \to B(S_1)$ respectively.

It only remains to prove Lemma 8. Here the ring object is $(C, N) = (\mathrm{End}\ TV, \mathrm{Hom}(TV, V))$, and $\rho : \mathcal{C}(N, N)$ is given by $\rho(f) = R_f$. To show that $N \otimes_C TV = V$, we need to :

1. define $h : TV \to \mathcal{C}(C, V)$
2. check that $h(ap) = h(p) \circ \rho(a)$ for all $p \in TV$, $a \in C$, and
3. check that $\tilde{h} : TV \to TV$ is an isomorphism.

Every $p \in TV$ gives $A(p) : k \to TV$ and induces therefore a $\mathcal{C}$-morphism $R_{A(P)}$ $\mathrm{Hom}(TV, V) = C \to \mathrm{Hom}(k, V) = V$. We define $h(p) = R_{A(p)}$.

This takes care of condition 1.

To see 2, we must show that $h(ap) = h(p) \circ \rho(a)$, for all $a \in \mathrm{End}(TV)$, for all $p \in TV$. The composite $k \xrightarrow{A(p)} TV \xrightarrow{a} TV$ is precisely $A(ap)$, so $h(ap) = R_{A(ap)} = R_{a \circ A(p)} = R_{A(p)} \circ R_a = h(p) \circ \rho(a)$.

Next we show that $\bar{\bar{h}}$ is the identity. Note that $T(h(p)) : \mathrm{End}\,(TV) \to TV$ is just $a \to a(p)$ for all $a \in \mathrm{End}\,(TV)$. Therefore $\bar{h}(p) = T(h(p))\,1 = p$. Q.E.D.

Proposition 3 : $f : A \to B$ is surjective if and only if $H(f) : |\,B\,| \to |\,A\,|$ is fully faithful and any exact sequence : $0 \to W' \to H(f)\,V \to W'' \to 0$ is isomorphic to the $H(f)$-image of an exact sequence

$$0 \to V' \to V \to V'' \to 0 \text{ in } |\,B\,|.$$

Proof : If f is indeed surjective, that these properties are enjoyed by $H(f)$ is absolutely clear.

Conversely, we have to show that $A \to B/J$ is surjective for all open two-sided ideals J of B given the hypothesis on $H(f)$. Let I be the kernel of $A \to B/J$. Then $0 \to A/I \to H(f)\,(B/J) \to H \to 0$ shows that there is $N \hookrightarrow B/J$ and a commutative diagram :

$$\begin{tikzcd} H(f)N \arrow[rr] \arrow[dr, "\cong"'] & & H(f)(B/J) \\ & A/I \arrow[ur] & \end{tikzcd}$$

In other words the image of A/I and B/J is a B-module, *i.e.*, it is an ideal on B/J. But 1 is in this image : Therefore $A/I \to B/J$ is an isomorphism. Q.E.D.

Remarks : 1. We have been purposely careless by identifying functors when in truth there is only a natural equivalence between them.

2. We could have extended the G of Lemma 1 to a $G : |\,A(\mathcal{C})\,| \to \mathcal{C}$ such that $F \circ G =$ identity. But this would have been much more tedious to write out. To check $F1$ and $F2$, defining G at a finite stage, *i.e.*, from $|\,A(S)\,| \to \mathcal{C}$ suffices as we have already seen.

§2. *Tannaka categories*

If $G = \mathrm{Spec}\,R$ is an affine group-scheme over k, let $|\,G\,|$ be the category of finite dimensional G-representations. If $A = R^*$, then A becomes an algebra in the sense of § 1 and it is easy to see that $|\,A\,| = |\,G\,|$ canonically. Let $T_k : |\,G\,| \to |\,k\,|$ be the "forgetful functor" as usual. The multiplication homomorphism $a \otimes b \mapsto ab$ from $R \otimes R \to R$ induces $\Delta : A \to A \hat{\otimes} A$. Given representations V and W of A, $V \otimes W$ becomes an $A \hat{\otimes} A$-module and by using Δ it becomes a A-module again. This is just the tensor product of representations V and W; it will be denoted by $V \hat{\otimes} W$. Let L_0 be the trivial representation.

Putting $|\,G\,| = \mathcal{C}$, $T_k = T$, the $(\mathcal{C}, T, \hat{\otimes}, L_0)$ has the following properties :

$\mathcal{C}1$, $\mathcal{C}2$, $\mathcal{C}3$: the pair $(\mathcal{C}, T)$ satisfies the hypothesis of Proposition 1,

$\mathcal{C}4$, $\hat{\otimes} : \mathcal{C} \times \mathcal{C} \to \mathcal{C}$ is a covariant functor which is k-linear in each variable, and

$$\begin{tikzcd} \mathcal{C} \times \mathcal{C} \arrow[r, "\hat{\otimes}"] \arrow[d, "T \times T"] & \mathcal{C} \arrow[d, "T"] \\ {} & {} \end{tikzcd}$$

commutes, where $\otimes$ is the usual tensoring functor.

$\mathcal{C}$ 5 : $\hat{\otimes}$ is associative preserving T : there is a natural equivalence of the functors from $\mathcal{C} \times \mathcal{C} \times \mathcal{C} \to \mathcal{C}$ given by $\hat{\otimes} \circ (l_{\mathcal{C}} \times \hat{\otimes})$ and $\hat{\otimes} \circ (\hat{\otimes} \times l_{\mathcal{C}})$ such that for all objects P, Q, R of $\mathcal{C}$, the isomorphism $H(P, Q, R)$ from $P \hat{\otimes} (Q \hat{\otimes} R)$ to $(P \hat{\otimes} Q) \hat{\otimes} R$ after an application of T gives the standard isomorphism $TP \otimes (TQ \otimes TR) \to (TP \otimes TQ) \otimes TR$ which gives the associativity of the tensor product for k-modules.

$\mathcal{C}$ 6 : $\hat{\otimes}$ is commutative preserving T (in the above sense).

$\mathcal{C}$7 : the functor $\mathcal{C} \to \mathcal{C}$ given by $P \mapsto L_0 \hat{\otimes} P$ is naturally equivalent to the identity functor, and there is an isomorphism $k \cong T_{L_0}$ so that for all $P \in \operatorname{Obj} \mathcal{C}$. $L_0 \hat{\otimes} P \xrightarrow{\cong} P$ yields after T an isomorphism :
$k \otimes TP \to TL_0 \otimes TP \to TP$ which is the standard isomorphism $a \otimes P \to ap$.

8 : If $L \in \operatorname{Obj} \mathcal{C}$ and TL has dimension one, there is a L^{-1} such that $L \hat{\otimes} L^{-1} \cong L_0$

$A(\mathcal{C}, T, \hat{\otimes}, L_0)$ satisfying all the above properties is called Tannaka category, The aim here is to prove :

Proposition 4 : Any Tannaka category is $|G|$ for a unique affine group-scheme G, and homomorphisms of Tannaka categories are induced by a homomorphism of affine group-schemes.

Proof: By Proposition 1, if $A = A(\mathcal{C})$, the pair $(\mathcal{C}, T)$ may be identified to $(|A|, T_k)$.

By Proposition 2, the axiom $\mathcal{C}$ 4 shows that $\hat{\otimes}$ is induced by a unique homomorphism $\triangle : A \to A \hat{\otimes} A$.

$\mathcal{C}$ 5 and $\mathcal{C}$ 6 show that

$(l_A \otimes \triangle) \circ \triangle = (\triangle \otimes l_A) \circ \triangle$ and $\triangle = \theta \circ \triangle$ where $\theta : A \hat{\otimes} A \to A \hat{\otimes} A$ is defied by $\theta(a \otimes b) = b \otimes a$.

$\mathcal{C}$ 7 shows that there is a homomorphism $\varphi : A \to k$ such that $A \xrightarrow{\triangle} A \hat{\otimes} A \xrightarrow{\varphi \otimes 1} k \otimes A = A$ is the identity.

Now let R be the vector space of continuous linear functionals on A, *i.e.* all linear functionals that vanish on some neighbourhood of zero.

$\triangle : A \to A \hat{\otimes} A$ gives a linear transformation $\triangle^* : R \otimes R \to R$. Then $\mathcal{C}$ 5 and $\mathcal{C}$ 6 show that $\triangle^*$ defines an associative commutative algebra-structure on R. And $\mathcal{C}$ 7 shows that this algebra R has an identity. Put $G = \operatorname{Spec} R$.

Now A is itself an algebra : thus $A \hat{\otimes} A \to A$ given by $a \otimes b \mapsto ab$ induces $\mu : R \to R \otimes R$. Because $\triangle : A \to A \hat{\otimes} A$ is an algebra homomorphism and not just a linear map, it follows that μ is a homomorphism of k-algebras. Thus μ induces $m : G \times G \to G$.

The associativity of the algebra structure on A shows that m makes G an affine semi-group-scheme (*i.e.*, the multiplication m is associative) and finally the identity of A gives an identity to G making G an affine monoid-scheme. We have to use $\mathcal{C}$ 8 to show that G is an affine group-scheme.

For a discrete monoid M there is a natural embedding $M \to k[M]$. This generalises in the above situation to a closed immersion $i : G \to \underline{A}$.

Any finite dimensional vector space V gives a scheme $\underline{V}$ by $\underline{V} = \operatorname{Spec} S(V^*)$. For any k-scheme X, $\Gamma(X, O_x) \otimes V = \operatorname{Mor}(X, \underline{V})$.

We define $\underline{A} = \operatorname{Spec} S(R)$. The reason being: for any k-scheme X, $\varinjlim_{I} \operatorname{Mor}(X, (\underline{A/I})) = \operatorname{Mor}(X, \underline{A})$ where the I run through open two-sided ideals of A.

The natural homomorphism $j : S(R) \to R$ given by $jx = x$ for all $x \in R$ induces a closed immersion $i : G \to \underline{A}$.

Next note that $\underline{A}$ is a monoid-scheme and in fact an inverse limit of the monoid-schemes $\underline{A/I}$. The operation $\underline{A} \times \underline{A} \to \underline{A}$ is given by $S(R) \to S(R) \otimes S(R)$ so that for $x \in R$ its image is $\mu \times \varepsilon R \otimes R \subseteq S(R) \otimes S(R)$. It is clear that $i : G \to \underline{A}$ is a homomorphism of monoid-scheme.

Similarly we form the schemes $(\underline{A/I})^*$ and $\underline{A}^*$; these are affine group-schemes for $f \in (A/I)$, $Nf =$ determinant of left-multiplication by f is a polynomial function on $(\underline{A/I})$, thus it is an element $d(I) \in S((A/I)^*)$. Put $(\underline{A/I})^* = \operatorname{Spec} \; S(A/I)^*)_{d(I)}$ and $\underline{A}^*$ is the spectrum of the ring got from inverting all the $d(I)$ in $S(R)$.

Let us now assume $\mathcal{C}$ 8. Then for an open two-sided ideal I, if $r = rk(A/I)$, consider $\Lambda^r(A/I)$. By $\mathcal{C}$ 8 it will follow that the composite $G \xrightarrow{i} \underline{A} \to \underline{A/I}$ has its image in $(\underline{A/I})^*$. This gives a factoring

$$\begin{tikzcd} G \arrow[r, "i"] \arrow[dr, "j"'] & \underline{A} \\ & \underline{A^*} \arrow[u] \end{tikzcd}$$

and clearly j is a closed immersion. That G is an affine group-scheme follows from

Lemma : If $G \to P$ is a closed monoid-scheme of an affine group-scheme, then G is an affine group-scheme.

Proof : Let $P = \operatorname{Spec} B$ and let I be the ideal defined by G. The morphism $Z : P \times P \to P \times P$ given by $(p_1, p_2) \mapsto (p_1, p_1 p_2)$ is an isomorphism. It suffices to show that Z induces an isomorphism from $G \times G$ to itself.

Let $Z^* : B \otimes B \to B \otimes B$ be the induced homomorphism on co-ordinate rings. Let $J = I \otimes B + B \otimes I$. Then $Z^*(J) \subseteq J$. We have to show that $Z^*(J) = J$.

It is well-known that B is the union of its finitely generated Hopf sub-algebras $\{C\}$. For any such C, Spec C is a group-scheme and Z^* restricts to an isomorphism of C. Thus if $(Z^*)^{-1}(J \cap (C \otimes C)) \neq J \cap (C \otimes C)$, then $(Z^*)^{-n}(J \cap (C \otimes C)) = J_n$ gives a strictly increasing sequence of ideals, which is not

possible because C is Noetherian. Therefore, $Z^*(J \cap (C \otimes C)) = J \cap (C \otimes C)$ for all C implying that $Z^*(J) = J$. This proves the lemma.

The second assertion about homomorphisms of Tannaka categories follows again from Proposition 2.

From now on $R(G)$ will denote the co-ordinate ring of an affine group-scheme G and $A(G)$ will denote its dual.

A homomorphism $G \to H$ is said to be surjective if $R(H) \to R(G)$ is injective ; equivalently if $A(G) \to A(H)$ is surjective.

Proposition 5 : A homomorphism $G \to H$ is surjective if and only if the corresponding functor $F : |H| \to |G|$ is fully faithful and for any exact sequence $0 \to W' \to FV \to W'' \to 0$ in $|G|$ there is an exact sequence $0 \to V' \to V \to V''$ $\to 0$ in $|H|$ and a commutative diagram :

$$\begin{tikzcd}
0 \arrow[r] & W' \arrow[r] \arrow[d, "="'] & FV \arrow[r] \arrow[d, "1"] & W'' \arrow[r] \arrow[d, "="] & 0 \\
0 \arrow[r] & FV' \arrow[r] & FV \arrow[r] & FW'' \arrow[r] & 0
\end{tikzcd}$$

This is an immediate consequence of Proposition 3.

Proposition 6 : An affine group-scheme is finite if and only if there is a finite set S of G-representations such that any representation of G is a sub-quotient of a finite direct sum of representations from S.

Proof : If G is finite, then any representation is contained in a direct sum of copies of $R(G)$ which is itself a finite-dimensional representation of G.

Conversely, given such a set, put $(\mathcal{C}, T) = (|G|, T_k)$. To prove that $R(G)$ is finite dimensional, it suffices to prove that $A(G) = A(\mathcal{C})$ is finite-dimensional. We shall show in fact that $A(\mathcal{C}) \to A(S)$ is an injection. Let $a \in A(\mathcal{C})$. Suppose that $\pi_V(a) = 0$ for all $V \in S$. If V and W belong to S, and p and q are the projection from $V \oplus W$ to V and W respectively, then $Tp \circ \pi_{V \oplus W}(a) = \pi_V(a) \circ Tp = 0$ and $Tq \circ \pi_{V \oplus W}(a) = \pi_W(a) \circ Tq = 0$. This shows that $\pi_{V \oplus W}(a) = 0$. Similarly $\pi_Q(a) = 0$, for all Q whenever Q is a finite direct sum of objects from S. If i : $P \to Q$ is an injection and $\pi_Q(a) = 0$, then $Ti \circ \pi_P(a) = \pi_Q(a) \circ Ti = 0$ showing that $\pi_P(a) = 0$. Similarly if $j : P \to Q$ is a surjection and $\pi_P(a) = 0$, then $\pi_Q(a) = 0$. Thus we have shown that $\pi_P(a) = 0$ for all sub-quotients of all finite direct sums of members of S, *i.e.*, $\pi_P(a) = 0 \ \forall\ P \in \operatorname{obj} \mathcal{C}$, and therefore $a = 0$. Q.E.D.

The last two propositions are just what are required to show that (see Chapter I, § 3) :

1. $\pi(X, x_0) \to \pi(S, x_0)$ is a surjection, for all finite sets S of essentially finite vector bundles.
2. $\pi(X, x_0) \to \varprojlim_S \pi(S, x_0)$ is an isomorphism with the S as above,
3. $\pi(S, x_0)$ is a finite group-scheme with the S as above.

References

Atiyah M F 1957 Vector bundles on an elliptic curve, *Proc. London Math. Soc.*, **Third Series, 7** 412–452

Grothendieck A 1965 *Elements de Geometrie Algebrique.*, I.H.E.S. Publications **24**

Kunze Ernst 1969 Characterizations of Regular local rings of characteristic *p*, *Am. J. Math.* **91** 772–784

Saavedra Rivano 1972 Categories Tannakiennes, Lecture Notes, Springer Verlag **265**

Safarevic I R 1956 On *p*-extensions, *Am. Math. Soc. Translations, Series* **2**, Vol. **4** 59–72

Seshadri C S 1967 Space of unitary vector bundles on a compact Riemann surface; *Ann. Math.* **85** 303–306

Seshadri C S 1977 Moduli of vector bundles on curves with parabolic structures, *Bull. Am. Math. Soc.* Vol. **83**, No. 1

Weil A 1938 Generalisation des fonctions abeliennes; *J. Mathematiques Pures et Appliques* **17** 47–87

Proc. Indian Acad. Sci. (Math. Sci.), Vol. 91, Number 2, July 1982, pp. 123–132.

Couette flow of a non-homogeneous fluid

K N VENKATASIVA MURTHY and K PONNURAJ
Department of Mathematics, Sri Krishnadevaraya University, Anantapur 515 003, India

MS received 20 March 1981 ; revised 30 October 1981

Abstract. The two dimensional Couette flow of a non-homogeneous viscous fluid is studied. The plane boundaries of the channel are maintained at different temperatures. The upper plane moves with a uniform horizontal velocity and the lower plane is at rest. The fluid is subjected to suction and injection at the boundaries. The steady equations are solved by introducing similarity variables which are expanded in series of powers of a small stratification parameter. The non-linear theory predicts that the temperature depends on the distance x from the throat section, an observation which is not predicted by the linear theory. The non-linear effects on velocity and temperature are studied. The rate of heat transfer is discussed.

Keywords. Stratification ; heat transfer ; suction ; injection.

1. Introduction

The flow of a viscous incompressible fluid between two plane boundaries is of fundamental interest in fluid dynamics. This problem is extensively used to find the interplay of various fluid forces like viscous force and buoyancy force. This configuration is used with suction and injection of fluid at the plane boundaries to solve some problems that arise in practice, like the phenomenon of separation [3; 6; 4]. This model is also used to discuss the heat transfer in acquiring the knowledge of several technological devices [1; 2]. The Couette flow is further useful in the theory of hydrodynamic lubrication. In view of these applications to modern technology the investigations on Couette flow remain important.

In this paper a two-dimensional Couette flow of a viscous stratified incompressible fluid, that arises in throat sections of channels is studied. The two plane boundaries of the channel are maintained at different temperatures. The fluid is injected at the upper plane and sucked at the lower plane with different velocities. The upper plane moves with a uniform velocity. The solution for the steady flow is obtained by assuming similarity variables for the velocities. The variables are expanded in suitable series of powers of a small stratification parameter. The formulation is based on the procedure adopted by Verma and Bansal [6] and Venkatasiva Murthy [5]. The velocity, temperature distributions and the rate of heat transfer are derived. They yield certain interesting features. The thick-

ness of the viscous and thermal boundary layers on the planes is determined for fluids of small kinematic viscosity. The non-linearity of the equations brings out certain properties of the temperature distribution which a linear theory cannot predict.

2. Formulation of the problem

The two-dimensional flow of a viscous stratified fluid between two infinite parallel planes $z' = \pm L/2$ is considered with reference to the rectangular cartesian coordinate axes $Ox'z'$. In the basic state the planes are maintained at different temperatures T_1 and T_2. The fluid is injected at the upper plane and sucked at the lower plane with a uniform velocity. The temperature difference between the planes causes an exponential density distribution in the fluid. If (u_e, v_e, w_e) are the velocity components, ρ_e the density and T_e the temperature in the basic state, then

$$u_e = v_e = 0, \quad w_e = - W'(\text{constant}),$$

$$T_e = T_0 + \frac{T_1 - T_2}{2}\left\{\frac{\cosh(\rho_0 c_p LW'/2K) - \exp(-\rho_0 c_p W' z'/K)}{\sinh(\rho_0 c_p LW'/2K)}\right\},$$

$$\rho_e - \rho_0 = -\frac{\epsilon\rho_0}{T_0}(T_e - T_0), \quad T_0 = (T_1 + T_2)/2.$$

The pressure p_e in the basic state satisfies the equation

$$\frac{\partial p_e}{\partial z} = -\rho_e g.$$

T_0 and ρ_0 are the mean values and ϵ is a small constant.

The disturbance in the fluid, over this basic state, is created when (i) the upper plane moves with velocity $\epsilon U'$ and (ii) the normal velocities at the permeables planes $z' = \pm L/2$ are maintained at $- W'(1 \pm \epsilon)$ respectively. The equations of motion for a steady flow are

$$\rho' \bar{q}' \cdot \nabla \bar{q}' = -\nabla \bar{p}' + \mu \nabla^2 \bar{q}' - \rho' g\bar{k},$$

$$\nabla \cdot (\rho' \bar{q}') = 0,$$

$$\rho' c_p \bar{q}' \cdot \nabla T' = K\nabla^2 T' + \phi,$$

$$\rho' - \rho_e = -\epsilon\frac{\rho_0}{T_0}(T' - T_e),$$

where ϕ is the viscous dissipation function and $\bar{k}$ is the unit vector in the z'-direction. The flow is two-dimensional and the velocity is $\bar{q}' = (u', O, w')$. The variables ρ', p', T' are the density, pressure and temperature respectively. The constant μ is the coefficient of viscosity. The boundary conditions are:

$$u' = \epsilon U', \quad w' = - W'(1 + \epsilon), \quad T' = T_1, \text{ when } z' = L/2;$$

$$u' = 0, \quad w' = - W'(1 - \epsilon), \quad T' = T_2, \text{ when } z' = - L/2.$$

Since the stratification parameter ϵ is small ($\epsilon \ll 1$) for all stable stratifications, we introduce the non-dimensional variables and expansions of variables in series of powers of ϵ as follows:

$$u' = \epsilon V u = \epsilon V (u_0 + \epsilon u_1 + \ldots), \tag{1}$$

$$w' = -W' + \epsilon V w = -W' + \epsilon V (w_0 + \epsilon w_1 + \ldots), \tag{2}$$

$$p' = p_e + \frac{\epsilon \rho_0 g L}{T_0}(T_1 - T_2) p, \tag{3}$$

$$p = p_0 + \epsilon p_1 + \epsilon^2 p_2 + \ldots, \tag{4}$$

$$T' = T_s + \epsilon (T_1 - T_2)\theta = T_s + \epsilon (T_1 - T_2)(\theta_0 + \epsilon \theta_1 + \ldots), \tag{5}$$

where

$$V = [(T_1 - T_2) g L / T_0]^{1/2}$$

is a constant, which is of the dimension of velocity. We also write $x' = Lx$, $z' = Lz$. With these non-dimensional variables and under the Boussinesq approximation that the influence of density variation with temperature is considered only on the body force term, the equations of motion in the non-dimensional form are

$$\epsilon u \frac{\partial u}{\partial x} + \epsilon w \frac{\partial u}{\partial z} - \delta \frac{\partial u}{\partial z} = -\frac{\partial p}{\partial x} + E \nabla^2 u, \tag{6}$$

$$\epsilon u \frac{\partial w}{\partial x} + \epsilon w \frac{\partial w}{\partial z} - \delta \frac{\partial w}{\partial z} = -\frac{\partial p}{\partial z} + E \nabla^2 w + \epsilon \theta, \tag{7}$$

$$\epsilon u \frac{\partial \theta}{\partial x} + \epsilon w \frac{\partial \theta}{\partial z} - \delta \frac{\partial \theta}{\partial z} + \frac{P\delta \exp(-P\delta z/E)}{2E \sinh(P\delta/2E)} w = \frac{E}{P} \nabla^2 \theta +$$
$$+ \epsilon H \left[\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial w}{\partial z}\right)^2 + \frac{1}{2}\left(\frac{\partial w}{\partial x} + \frac{\partial u}{\partial z}\right)^2 \right]. \tag{8}$$

$$\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} = 0, \tag{9}$$

where $E = \mu/\rho_0 VL$ is the inverse of a Reynolds number, $P = \mu c_p / K$ is the Prandtl number, $\delta = W'/V$ is suction parameter and $H = 2\mu V / L\rho_0 c_p (T_1 - T_2)$.

In view of the expansions in (1) to (5) we obtain the following differential equations:

$$E \nabla^2 u_0 - \frac{\partial p_0}{\partial x} + \delta \frac{\partial u_0}{\partial z} = 0, \tag{10}$$

$$E \nabla^2 w_0 - \frac{\partial p_0}{\partial z} + \delta \frac{\partial w_0}{\partial z} = 0, \tag{11}$$

$$u_0 \frac{\partial u_0}{\partial x} + w_0 \frac{\partial u_0}{\partial z} - \delta \frac{\partial u_1}{\partial z} = -\frac{\partial p_1}{\partial x} + E \nabla^2 u_1, \tag{12}$$

$$u_0 \frac{\partial w_0}{\partial x} + w_0 \frac{\partial w_0}{\partial z} - \delta \frac{\partial w_1}{\partial z} = -\frac{\partial p_1}{\partial z} + E \nabla^2 w_1 + \theta_0. \tag{13}$$

The energy equation gives

$$\frac{E}{P}\nabla^2\theta_0 + \delta\frac{\partial\theta_0}{\partial z} = \frac{P\delta\exp(-P\delta z/E)}{2E\sinh(P\delta/2E)}w_0, \tag{14}$$

$$\frac{E}{P}\nabla^2\theta_1 + \delta\frac{\partial\theta_1}{\partial z} = \frac{P\delta\exp(-P\delta z/E)}{2E\sinh(P\delta/2E)}w_1 + u_0\frac{\partial\theta_0}{\partial x} + w_0\frac{\partial\theta_0}{\partial z}$$

$$-H\left\{\left(\frac{\partial u_0}{\partial x}\right)^2 + \left(\frac{\partial w_0}{\partial z}\right)^2 + \frac{1}{2}\left(\frac{\partial w_0}{\partial x} + \frac{\partial u_0}{\partial z}\right)^2\right\}\cdots \tag{15}$$

The continuity equation suggests similarity variables

$$u(x,z) = x\frac{\partial f}{\partial z} + g,\ w(z) = -f(z),$$

and hence

$$u_i(x,z) = x\frac{\partial f_i}{\partial z} + g_i,\ w_i(z) = -f_i(z),\ i = 0, 1, 2, \ldots,$$

where f_i and g_i are unknown functions of z. We eliminate pressure from (6) and (7) to obtain the general form for θ as

$$\theta(x,z) = x^2\theta^{(2)}(z) + x\theta^{(1)}(z) + \theta^{(0)}(z),$$

so that, for $\theta = 0, 1, 2, \ldots,$

$$\theta_i(x,z) = x^2\theta_{i2}(z) + x\theta_{i1}(z) + \theta_{i0}. \tag{16}$$

Equations (6) and (7) then determine pressure as

$$p = x^2 p^{(2)}(z) + xp^{(1)}(z) + p^{(0)}(z) + Ax^2 + Bx + C,$$

where A, B, C are arbitrary constants of integration.
By choosing the origin such that $B = 0$, $A \neq 0$ [6; 5], we obtain the form for p as

$$p = x^2 p^{(2)}(z) + xp^{(1)}(z) + p^{(0)}(z) + Ax^2 + C,$$

and hence the form for p_i $(i = 0, 1, 2 \ldots,)$ as

$$p_i = x^2 p_{i2}(z) + xp_{i1}(z) + p_{i0}(z) + A_i x^2 + C_i, \tag{17}$$

where p_{i0}, p_{i1}, p_{i2} are unknown functions and A_i, C_i are arbitrary constants. The boundary conditions are

$$f_0 = \delta,\ f_1 = 0,\ f_0' = f_1' = 0,\ g_0 = U,\ g_1 = 0 \text{ on } z = \tfrac{1}{2}, \tag{18a}$$

$$f_0 = -\delta,\ f_1 = 0,\ f_0' = f_1' = 0,\ g_0 = 0,\ g_1 = 0 \text{ on } z = -\tfrac{1}{2} \tag{18b}$$

$$\theta_{i2} = \theta_{i1} = \theta_{i0} = 0 \text{ at } z = \pm\tfrac{1}{2}, \tag{18c}$$

where $U = U'/V$. It can be shown from (14) and (18c) that

$$\theta_{02} = \theta_{01} = 0. \tag{19}$$

Equations (16), (17) and (19) are substituted in equations (10) to (13) to obtain the following differential equations for f_0, g_0, f_1 and g_1.

$$f_0^{iv} + (\delta/E) f_0''' = 0, \tag{20}$$

$$g_0'' + (\delta/E) g_0' = 0, \tag{21}$$

$$E f_1^{iv} + \delta f_1''' - f_0' f_0'' + f_0 f_0''' = 0, \tag{22}$$

$$E g_1'' + \delta g_1' - g_0 f_0' + f_0 g_0' = 0. \tag{23}$$

The solutions of (20)–(23) subject to boundary conditions (18a, b) are

$$f_0 = A + Bz + Cz^2 + D\exp(2Nz),$$

$$g_0 = F + G\exp(2Nz),$$

$$f_1 = A_1 + A_2 z + A_3 z^2 + A_4 z^3 + A_5 z^4 + A_6 \exp(2Nz) + + z(A_7 + A_8 z + A_9 z^2)\exp(2Nz),$$

$$g_1 = B_1 + B_2 z + B_3 z^2 + (B_4 + B_5 z + B_6 z^2 + B_7 z^3)\exp(2Nz).$$

Also, from (14), (15) and (18c) we obtain

$$\theta_0 = X_0 + X_1 \exp(-P\delta z/E) + \psi(z),$$

$$\psi(z) = \frac{\exp(-P\delta z/E)}{2\sinh(P\delta/2E)} \left[\frac{CP}{3E} z^3 + \left(\frac{BP}{2E} + \frac{C}{\delta}\right) z^2 + + \left(\frac{2CE}{P\delta^2} + \frac{B}{\delta} + \frac{AP}{E}\right) z - \frac{P^2 D\delta \exp(2Nz)}{2NE(2NE - P\delta)}\right],$$

$$\theta_{10} = L_0 + L_1 \exp(-P\delta z/E) + \psi_1(z),$$

$$\theta_{11} = M_0 + M_1 \exp(-P\delta z/E) + Y_0 \exp(2Nz) + Y_1 \exp(4Nz),$$

$$\theta_{12} = C_0 + C_1 \exp(-P\delta z/E) + Y_2 z + Y_3 \exp(2Nz) + Y_4 \exp(4Nz),$$

$$\psi_1(z) = (d_0 z + d_1 z^2 + d_2 z^3) + (d_3 + d_4 z)\exp(2Nz) + d_5 \exp(4Nz)$$
$$+ (d_6 z + d_7 z^2 + d_8 z^3 + d_9 z^4 + d_{10} z^5 + d_{11} z^6)\exp(-P\delta z/E)$$
$$+ (d_{12} + d_{13} z + d_{14} z^2 + d_{15} z^3)\exp\left(2N - \frac{P\delta}{E}\right) z$$
$$+ d_{16} \exp\left(4N - \frac{P\delta}{E}\right) z + d_{17} \exp(-2P\delta z/E),$$

where

$$N = -\delta/2E,\ A = R\delta\left(\cosh N - \frac{N}{2}\sinh N\right),$$

$$B = 2\delta NR\cosh N,\ C = B\tanh N,\ D = -R\delta,\ F = -\frac{U\exp(-N)}{2\sinh N},$$

$$G = U/(2\sinh N),\ R = (N\cosh N - \sinh N)^{-1},$$

$$X_1 = \left[\psi\left(\frac{1}{2}\right) - \psi\left(-\frac{1}{2}\right)\right] \Big/ 2\sinh\,(P\delta/2E),$$

$$X_0 = -\psi\left(\frac{1}{2}\right) - X_1 \exp\,(-P\delta/2E),$$

and all other constants appearing in the above solution are lengthy expressions which depend on P, δ, E and H and will run for pages together if presented. Hence they are not presented to save space.

3. Discussion of the results

The velocity and temperature distributions are derived to order ϵ^2. The expressions for the velocities u and w consist of the exponentially decaying terms in z which decay at distances of order (E/δ). When E is small in comparison with δ, the exponential terms represent a boundary layer flow (at the upper plane if $\delta > 0$ and at the lower plane if $\delta < 0$) outside which the horizontal velocity distribution is linear to order ϵ. The differential equations for f_2 and g_2 would depend on θ while those for f_0, g_0, f_1 and g_1 do not depend on the temperature. Hence the temperature produces only changes of order ϵ^2 in the velocities. Since θ_1 contains an exponential term $\exp\,(-P\delta z/E)$, the temperature variations bring into effect additional exponential modes in the solution representing viscous boundary layers, of amplitude order ϵ^2 and thickness order $E/P\delta$ when $E \ll P\delta$.

The exponential terms in the solution for temperature tend to zero at distances of order E/δ, $E/P\delta$, $E/(P+1)\,\delta$, $E/(P+2)\,\delta$. When E is small in comparison with P and δ, these exponential terms represent thermal boundary layers of small thickness near the boundaries. While θ_0 does not depend on x, θ_1 is a function of both x and z. The temperature distribution, to order ϵ^2, is parabolic in x and even though ϵ is small, $\theta_1\,(x, z)$ can give a non-negligible contribution to the temperature at larger distances x. It may also be noted that this parabolic dependence on x is a non-linear effect of the viscous dissipation.

The velocity and temperature distributions are plotted in figures 1–8. The values of E and δ are chosen to have a greater order than ϵ to make the expansions assumed in the theory valid. Figures 1 and 2 show that as E increases the velocity decreases whereas the velocity increases as δ increases. Figure 3 shows the velocity distribution at various values of x. The continuous lines show the prediction of the linear theory and the dotted lines give the non-linear theory. When E is small in comparison with δ the steep rise in the velocity near the lower plane shows the boundary layer type behaviour of the velocity as observed earlier. It can be seen that the non-linear theory predicts a smaller velocity than the linear theory. Even though the non-linear contribution is small for smaller values of x, it increases with x and hence the linear theory cannot be taken to represent the velocity at moderate values of x, at least in the parts of the channel away from the two plane boundaries.

The velocity for negative values of W' (negative values of δ) is plotted in figure 4. The fluid is now injected at the lower plane and sucked at the upper plane. The velocity now decreases first and then increases to its value at the

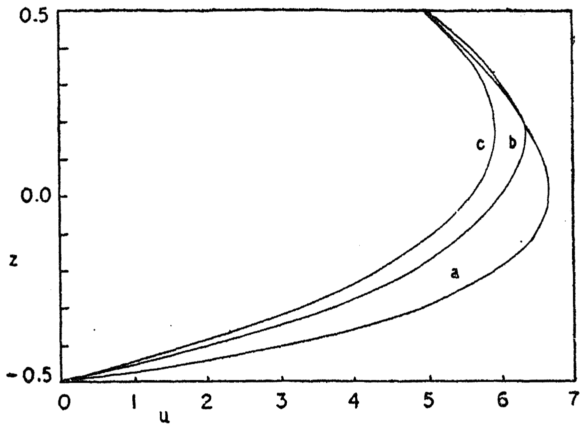


Figure 1. Velocity distribution when $\delta = 0{\cdot}05$, $U = 5$, $x = 20$. (a) $E = 0{\cdot}02$, (b) $E = 0{\cdot}05$, (c) $E = 0{\cdot}1$.

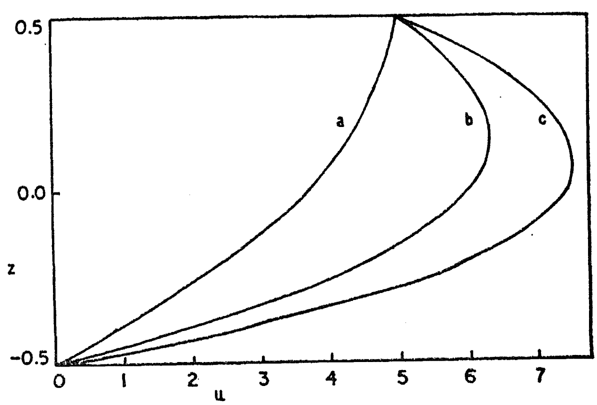


Figure 2. Velocity distribution when $E = 0{\cdot}05$, $x = 20$. (a) $\delta = 0{\cdot}02$, (b) $\delta = 0{\cdot}05$ (c) $\delta = 0{\cdot}07$

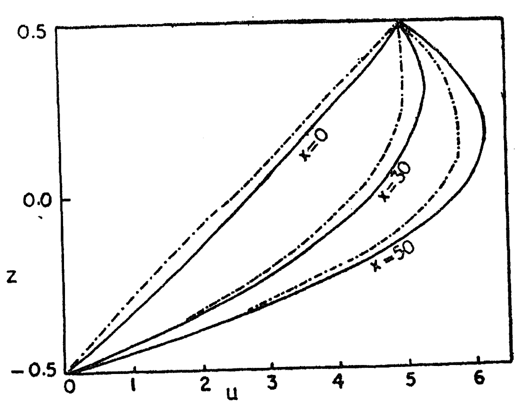


Figure 3. Velocity distribution when $\delta = 0{\cdot}02$, $E = 0{\cdot}05$. ———— linear theory, - · - · - · - non-linear theory.

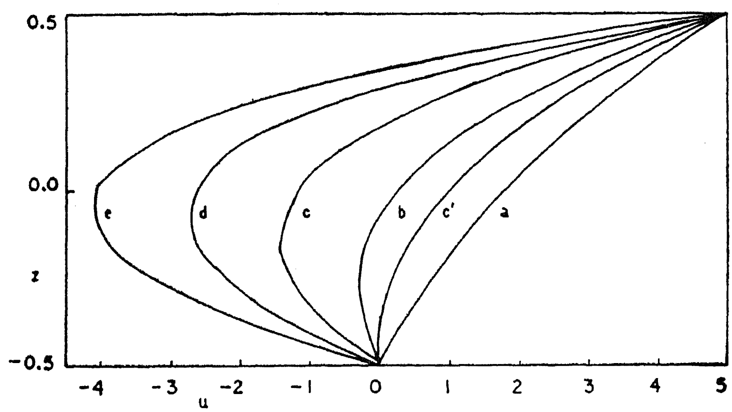


Figure 4. Velocity distribution when δ is negative, $\delta = -0{\cdot}05$. $E = 0{\cdot}05$, (a) $x = 0$, (b) $x = 10$, (c) $x = 20$, (d) $x = 30$, (e) $x = 40$, (c′) $\delta = -0{\cdot}02$, $E = 0{\cdot}05$, $x = 20$.

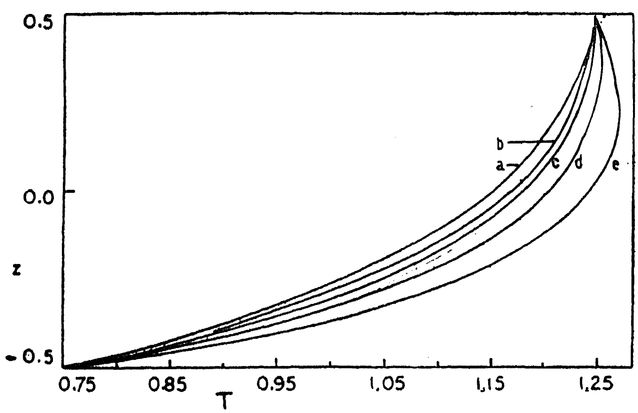


Figure 5. Temperature distribution when $P = 3{\cdot}5$, $\delta = E = 0{\cdot}05$. (a) $x = 10$, (b) $x = 40$, (c) $x = 60$, (d) $x = 80$, (e) $x = 100$.

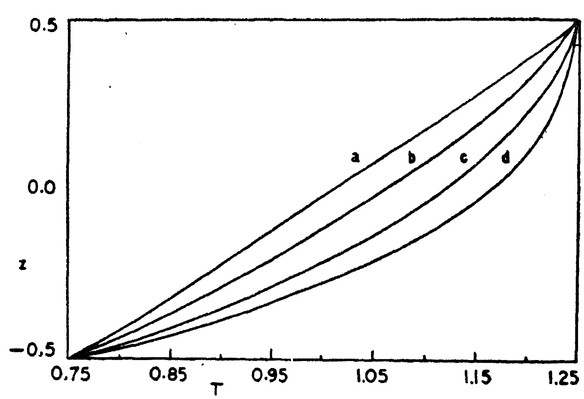


Figure 6. Temperature distribution when $x = 40$, $\delta = E = 0{\cdot}05$. (a) $P = 0{\cdot}5$, (b) $P = 1{\cdot}5$, (c) $P = 2{\cdot}5$, (d) $P = 3{\cdot}5$.

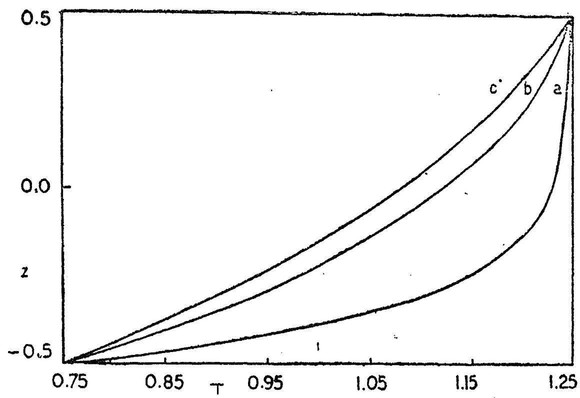


Figure 7. Temperature distribution when $P = 2{\cdot}5$, $\delta = 0{\cdot}05$, $x = 10$. (a) $E = 0{\cdot}02$ (b) $E = 0{\cdot}05$, (c) $E = 0{\cdot}1$.

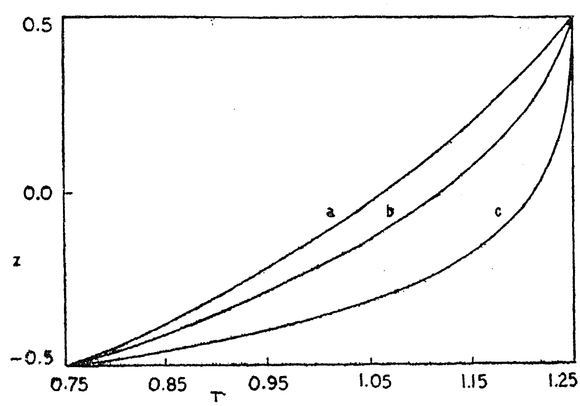


Figure 8. Temperature distribution when $P = 2{\cdot}5$, $E = 0{\cdot}05$, $x = 10$. (a) $\delta = 0{\cdot}02$, (b) $\delta = 0{\cdot}05$, (c) $\delta = 0{\cdot}1$.

upper plane and the velocity curves are comparable to those described by Verma and Bansal [6]. The flow from large values of x to the mouth of the channel is developed near the stationary wall. Thus the dragging action of the faster layers exerted on the fluid particles in the neighbourhood of the stationary wall is insufficient to overcome the influence of the adverse pressure gradient that develops. The velocity decreases (to larger negative values) with x increasing. Also as δ increases in value the velocity increases.

The temperature distribution at various cross-sections of the channel is shown in figure 5. The temperature depends on x parabolically to order ϵ^2. This increases the temperature in the interior beyond its value at the upper plane, as x increases. The important property that the temperature depends on x is brought out by the non-linear theory. The figures 6, 7 and 8 describe the behaviour of the temperature to the changes in the parameters P, δ and E. As P increases there is a rapid change in the temperature at the lower plane as the heat transfer will be more effective for larger P. As E increases the temperature decreases and as δ increases the temperature increases. There is a rapid change

in the temperature in a layer near the lower plane and then it remains almost a constant in the upper half of the channel when E is relatively smaller than P and δ.

The rate of heat transfer at the planes is

$$q' = K\left(\frac{(\partial T')}{\partial z'}\right)_{z' = \pm L/2}.$$

Defining the non-dimensional rate of heat transfer as $q = -q' L/K(T_1 - T_2)$ we obtain the rate of heat transfer at $z = \pm \frac{1}{2}$, to order one, given by

$$q = \frac{P\delta \exp(\mp P\delta/2E)}{2E \sinh(P\delta/2E)}.$$

We will discuss the rate of heat transfer to order one. It is found that perturbation of order ϵ will not altogether change the general behaviour of the rate of heat transfer. The rate of heat transfer at $z = \frac{1}{2}$ is

$$q_1 = \frac{2X}{\exp(2X) - 1}$$

and the rate of heat transfer at the lower plane is

$$q_2 = \frac{2X}{1 - \exp(-2X)}$$

where $X = P\delta/2E$. This shows that as X increases q_1 decreases and q_2 increases. Thus as E decreases or as P increases the rate of heat transfer decreases at the upper plane and increases at the lower plane. Also as $P\delta/2E$ decreases to zero q_1 and q_2 will tend to unity.

Acknowledgement

The authors are thankful to the referee for suggestions to improve the manuscript. One of the authors, K Ponnuraj is thankful to CSIR for the award of a junior research fellowship.

References

[1] Bleviss Z D 1958 *J. Aeronaut. Sci.* **25** 601

[2] Leadon B M 1957 *Convair Sci. Res. Lab.* RN 13

[3] Sinha K D and Choudhary R C 1965 *Proc. Indian Acad. Sci.* **A61** 308

[4] Terril R M and Sreshta G M 1965 *Z. Angew. Math. Phys.* **16** 470

[5] Venkatasiva Murthy K N 1980 *Proc. Indian Acad. Sci. (Math. Sci.)* **89** 103

[6] Verma P D and Bansal J L 1966 *Proc. Indian Acad. Sci.* **A64** 385

Proc. Indian Acad. Sci. (Math. Sci.), Vol. 91, Number 2, *July* 1982, pp. 133–138.

Arithmetic lattices in semisimple groups*

M S RAGHUNATHAN
School of Mathematics, Tata Institute of Fundamental Research, Bombay 400 005 India

MS received 26 June 1981

1. Introduction

Borel [1] showed that given a (connected) real semisimple Lie group G, it admits a discrete (arithmetic) subgroup Γ such that G/Γ is compact. In this paper we will establish the following refinement of that theorem.

Theorem. Let G be a connected linear semisimple Lie group and A a commutative group consisting of involutive automorphisms of G. Then G admits a discrete (arithmetic) subgroup Γ such that $G^a/\Gamma \cap G^a$ is compact for each $a \in \tilde{A}$, G^a being the fixed point set of a in G and $\tilde{A}$ is an abelian group of involutive automorphisms of G containing A and a cartan involution of G.

As was the case with Borel's proof, the theorem can be deduced from a result on Lie algebras. We omit the details of this deduction.

Theorem. Let g be a semisimple Lie algebra and A a commutative group consisting of involutive automorphisms of g. Then there is a Q-structure on g such that all elements of A are Q-rational and g admits a cartan involution defined over Q and commuting with A.

The kind of Q-structure introduced on g in the special case when g is compact has the additional property that all representations of g defined over R are equivalent to representations defined over Q.

The refined version proved here is likely to be of some interest in the context of geometric constructions for homology of compact locally symmetric spaces given by Millson–Raghunathan [4] and Millson [1]; in the special case where A is trivial, we get Borel's theorem.

Borel's theorem was preceded by results in the case of many classical groups. Siegel [5] initiated the subject by making the first constructions of uniform arithmetic subgroups in classical groups beyond $SL(2, R)$. This was generalised to cover a wilder class of classical groups by Klingen [2]. Ramanathan [3] pointed

* To Prof. K G Ramanathan on his 60th birthday.

out further examples and raised the question (in oral conversations) whether any semisimple Lie groups admits a uniform lattice.

2. The standard Q-form of a compact Lie algebra

Let k be a compact semisimple Lie algebra and $\boldsymbol{k} = k \otimes_R C$. Let $t \subset k$ be a cartan subalgebra and $\boldsymbol{t}$ = C-span of t. Let Φ be the root system of $\boldsymbol{k}$ with respect to $\boldsymbol{t}$ and for $\alpha \in \Phi$, let $\boldsymbol{k}(\alpha)$ denote the root space of α. As is well-known there exists a Chevalley basis of $\boldsymbol{k}$ viz., we have $\{H_\phi / \varphi \in \Phi\} \subset it$ and $E(\varphi) \in \boldsymbol{k}(\varphi)$, $\varphi \in \Phi$ such that

(i) $[H_\phi, E(\psi)] = 2\langle\varphi, \psi\rangle / \langle\psi, \psi\rangle \cdot E(\psi)$

(ii) $[H(\varphi), E(\psi)] = N_{\phi,\psi} E_{\phi+\psi}$ with $N_{\phi,\psi} \in Z$, $\varphi + \psi \in \Phi$

(iii) $[E(\varphi), E(-\varphi)] = H\varphi$.

The complex conjugation in $\boldsymbol{k}$ takes each $\boldsymbol{k}(\varphi)$ into $\boldsymbol{k}(-\varphi)$ so that for $\varphi \in \Phi$, $\bar{E}(\varphi) = \lambda(\varphi)$. $E(-\varphi)$ for some $\lambda(\varphi) \in C^*$. Since $\langle E(\varphi), \bar{E}(\varphi)\rangle > 0$, we conclude that $\lambda(\varphi) > 0$. Let $x \in \boldsymbol{T}$ the adjoint torus of $\boldsymbol{t}$ be chosen such that $\alpha(x) = \lambda(\alpha)^{-1/2} > 0$ for $\alpha \in \triangle$, a simple system of roots of k. If we set $E'(\varphi) = \lambda(\varphi)^{-1/2} E(\varphi) = Ad\, x\, E(\varphi)$, we see that for simple $\varphi \in \triangle$, $\bar{E}'(\varphi) = \lambda(\varphi)^{1/2} E(-\varphi) = E'(-\varphi)$ so that the complex conjugation takes $E'(\varphi)$ into $E'(-\varphi)$ for all $\varphi \in \triangle$. It follows immediately that $\bar{E}'(\varphi) = \pm E(-\varphi)$ for all $\varphi \in \Phi$ as well. The $E'(\varphi)$, $\varphi \in \Phi$ together with the $\{H_\alpha \mid \alpha \in \triangle\}$ constitute again a Chevalley basis. Let $\boldsymbol{k}_0$ be the $Q(i)$-span of the $\{E'(\varphi) \mid \varphi \in \Phi\}$ and the $\{H_\alpha \mid \alpha \in \triangle\}$. Then $\boldsymbol{k}_0$ is a $Q(i)$—split form of $\boldsymbol{k}$. Let k_0 be the fixed points in $\boldsymbol{k}_0$ of the complex conjugation: this is an antilinear involution over $Q(i)$. Then k_0 is a Q-form of $\boldsymbol{k}$. For each $\varphi > 0$, it is easily seen that the Lie algebra $\boldsymbol{a}_0(\varphi)$ spanned by $E'(\pm\varphi)$ and $H(\varphi)$ over $Q(i)$ is $Q(i)$—isomorphic to $SL(2)$, is stable under the conjugation with fixed algebra $a_0(\varphi)$ isomorphic over Q to $SU(2)$ the standard special unitary group over $Q(i)$ given by the hermitian form $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. An immediate consequence is that the reflection s_ϕ corresponding to φ in the Weyl group W of the adjoint algebraic group $\boldsymbol{K}$ with $\boldsymbol{k}$ as Lie algebra has a Q-rational representative in $N(\boldsymbol{T})$ the normaliser of $\boldsymbol{T}$ in $\boldsymbol{K}$ (for the natural Q-structure on k defined above).

In particular the unique element $w_0 \in W$ which takes all of $\triangle$ into negative roots has a Q-rational representative $w_0 \in N(\boldsymbol{T})(Q)$. Let $\boldsymbol{S}$ be the identity component of the group $\{x \in \boldsymbol{T} \mid w_0 x w_0^{-1} = x\}$. Then on T/S, w_0 acts by $w_0(x) = x^{-1}$. Further in $N(T)/S$ we have $w_0^2 \in T/S$ so that $w_0 w_0^2 w_0^{-1} = w_0^{-2} = w_0^2$ leading to the conclusion that w_0^2 is an element of order 2 modulo S. Note that $\boldsymbol{S}$ is defined over Q.

Definition. The Q-structure defined above will be called a Standard Q structure on the pair $(\boldsymbol{K}, \boldsymbol{T})$.

Proposition. Let $\boldsymbol{G}$ be a Q-algebraic group such that the identity component $\boldsymbol{G}^0$ of $\boldsymbol{G}$ is a torus and $\boldsymbol{G}/\boldsymbol{G}^0$ is abelian with every element of order 2. Suppose that $\boldsymbol{G}(Q) \to (\boldsymbol{G}/\boldsymbol{G}^0)(Q) = \boldsymbol{G}/\boldsymbol{G}_0$ is onto and the sequence.

$$(*)\ 1 \to \boldsymbol{G}^0 \to \boldsymbol{G} \to \boldsymbol{G}/\boldsymbol{G}^0 \to 1$$

admits a splitting γ over R and that the torus G^0 is anisotropic over R and splits over $Q(i)$. Then (*) splits over Q as well and the Q-splitting can be chosen to be conjugate to γ by an element of $G^0(R)$.

Proof. We argue by induction on dim G. We note first that every subtorus of G^0 defined over R is automatically defined over Q. Let $X(G^0)$ be the abelian group of 1 parameter subgroups of G^0. The Galois group Gal $(Q(i)/Q) \simeq$ Gal (C/R) operates on this by $\chi \to -\chi$. The group G/G^0 acts on $X(G^0)$ as well and has an eigen vector in $X(G^0) \underset{Z}{\otimes} Q$ hence in $X(G^0)$. Let S denote the corresponding torus in T. S is evidently defined over Q and normal in G. Let $G' = G/S$. Then by induction hypothesis we can find $u \in G^0(R)$ such that $\bar{\rho} = \bar{u}(\pi \circ r)\bar{u}^{-1}$ is defined over Q where $r : G/G^0 \to G$ is the given R-splitting, and $\pi : G \to G/S$ is the natural map and $\bar{u} = \pi(u)$. If we now set $H = \pi^{-1}(\rho(G/G^0))$, H is defined over Q and its identity component $H^0 = S$. We are thus reduced to the case when dim $G = 1$. First consider the action of the group G/G^0 on G_0. Since dim $G = 1$, the automorphism group of G is of order 2; it follows that G/G^0 has a subgroup B of index almost 2 which acts trivially on G^0. If $p : G \to G/G_0$ is the natural map $p^{-1}(B)$ is abelian—note that we have a splitting over R- and hence diagonalisable. Now we have the exact sequence

$$O \to X^*(B) \to X^*(p^{-1}(B)) \to X^*(G^0) \to 0$$

of the character groups. These are modules over Gal $(C/R) \cong$ Gal $Q(i)/R$ and by assumption the sequence is split as modules over Gal C/R hence also over Gal $Q(i)/Q$. Moreover *any* R-splitting is a Q-splitting ($X(B)$ is a trivial Galois-module). Thus we conclude that $p^{-1}(B)$ admits a Q-splitting of the form $B \cdot G^0$. The character group is a direct sum $X^*(B) \oplus X^*(G^0)$ with the action G/G^0 trivial on $X(B)$ and nontrivial on $X(G^0) \simeq Z$; if $B \neq G/G^0$, $X(B)$ then can be characterised as those elements which are fixed by G^0 as well as G/G^0. It is immediate now that B is normal in G. Consider then the quotient $H = G/B$. H^0 is isomorphic to G^0 and is hence 1-dimensional. The sequence

$$** \qquad 1 \to H^0 \to H \xrightarrow{q} H/H_0 \to Z/2 \to 1$$

is assumed to be split over R. Let $\tau \in H/H^0$ be the nontrivial element. Then $q^{-1}(\tau)$ is a principal homogeneous space over Q; it has a rational point over Q by assumption ($G(Q) \to G/G^0$ was assumed surjective). Now let τ_0 be the lift of τ given by the splitting over R and τ_0' a lift over Q. Then we have $\tau_0' = \tau_0 \cdot x$, $x \in H^0(R)$ so that

$$(\tau_0')^2 = \tau_0 x \cdot \tau_0 x = \tau_0^2 = 1$$

Thus τ_0' also gives a splitting of (**); in order to assert that τ_0' is a conjugate of τ_0 we need only have that x is a square of an element y in $H^0(R)$: for then

$$y^{-1}\tau_0 y = \tau_0 \cdot \tau_0^{-1} y^{-1} \tau_0 y = \tau_0 \cdot x.$$

Now $H^0(R)$ is isomorphic to the circle group, hence each $x \in H^0(R)$ is a square.

We obtain the required Q-splitting by taking the inverse image under $f: G \to G/B$ of the group $(\tau_0', 1)$. This completes the proof of the proposition.

Corollary. Let K be a compact (connected semisimple) group and $A \subset \text{Aut } K$ be an abelian subgroup consisting entirely of elements of order 2. Then there is a A-stable torus T in K and a "standard" Q-structure on (k, t) with A consisting entirely of Q-rational automorphisms of k.

Proof. We assume $K = (\text{Aut } K)^0$. We fix a maximal subgroup A_1 of A which is contained in some maximal torus. Let $z(A_1)$ be the fixed point set of A_1 in k. Then $z(A_1)$ is A-stable. Moreover a maximal abelian subalgebra of $z(A_1)$ is maximal abelian in k as well. Since A consists of elements of order 2, A has a common eigen vector $X \in k$. The corresponding torus in K is evidently A-stable. Hence there is among abelian subalgebras of $z(A_1)$, a maximal non zero one say b which is A-stable. Since b is A-stable so is $z_1(b)$ the centraliser of b in $z(A_1)$. If b is not maximal abelian its orthogonal complement in $z_1(b)$ will contain a 1-dimensional A-stable subspace leading to a contradiction. Thus b is a A-stable cartan subalgebra of k. We denote the corresponding torus by T. Take now any standard Q-structure on (k, t). The group A is a direct product $A_2 \times A_1$ where $A_2 \cap T = \{1\}$ and $A_1 \subset T$. A_1 consists of elements of order 2 and these are easily seen to be Q-rational. By Proposition we can find $x \in T(R)$ which conjugates A_2 into Q-rational points. Replacing the Chevalley basis we started out with for defining the standard structure by their transforms under $Ad\, x^{-1}$ we obtain all the requisite properties. Observe that as $x \in T(R)$ the Q-structure on T remains unchanged. The Q-structure on k remains isomorphic to the original one as well as is easily seen. If $N(T)$ = normaliser T in $\text{Aut}(k)$, $N(T)(Q) \xrightarrow{\pi} N(T)/T = [N(T)/T](Q)$ gives surjection at the Q-rational level as the Dynkin automorphisms fixing T is also Q-rational (all the hypothesis of the proposition are satisfied by $G = \pi^{-1}\pi(A)$ and $G^0 = T$).

Lemma. Let G be a connected linear semisimple Lie group and $A \subset \text{Aut } G$ a finite abelian group consisting of involutions. Then G admits a cartan involution commuting with A.

Proof. Let K be a maximal compact subgroup of Aut G containing A. K defines a cartan involution of G which evidently commutes with all the elements of A.

Theorem. Let G be a connected linear semisimple Lie group and g its Lie algebra. Let $A \subset \text{Aut } G$ be any group of commuting involutions of G. Then g admits a Q-structure such that all $a \in A$ are Q-rational and there is a Q-rational cartan-involution commuting with A as well.

Proof. Enlarge A to include a cartan involution θ (*cf.* Lemma above). Let $g = u + p$ be the cartan-decomposition with u compact. Then u and p are A-stable as all of A commute with θ. Let $k = u + ip$. Then k is a compact Lie algebra. By proposition we can find a A-stable torus t in k such that (k, t) admits a standard Q-structure with $A \subset K(Q)$. Since θ is Q-rational u and ip are defined over Q for this Q-structure. This immediately gives a Q-structure on $u + p = g$ as well. Next since each $a \in A$ acts Q-rationally on u as well as ip and hence on p, each $a \in A$ is Q-rational for this Q-structure on g.

3. Representations of the standard Q-form

The following property of the standard Q-form of k seems to be of some interest.

Theorem. Let k_Q be a standard Q-form of (k, ι) with k a compact semisimple Lie algebra. Then every representation of k_Q defined over R is equivalent to a unique one defined over Q.

In view of complete reducibility, it suffices to show that each irreducible R-representation of k_Q is equivalent to one defined over Q. (The uniqueness part of the statement is easy to prove: one way is to use the Zariski density of $K(Q)$ in K (K = simply connected Q algebraic group determined by k_Q) and use the fact that representations of $K(Q)$ are characterised by their characters: see for instance Van der Wearden [6, exercise, p. 175]. In fact it suffices to show that each irreducible representation over Q of k_Q remains irreducible over R. To see this observe that if σ is an irreducible representation of k_Q defined over R, σ may be assumed to be defined over some number field; the set of all representations of k_Q on a fixed finite dimensional vector space is a variety V defined over Q and $\sigma \in V(R)$. The orbit of σ under K is open in $V(R)$ in view of the Whitehead lemma and hence contains $\bar{Q}$-rational points. We may thus assume σ to be defined over a real number field $L \supset Q$, with L of minimal possible degree. Consider now the underlying L vector space as a Q vector space and denote the corresponding representation by τ. Since L commutes with the action of $K(Q)$ and L-span of any non zero $K(Q)$-irreducible Q-subspace of $W(\sigma)$ (= representation space of σ) is all of $W(\sigma)$, we conclude that τ is isotypical of fixed type τ_0. Evidently, $W(\sigma)$ is a quotient of $W(\tau_0) \otimes_Q L$. Since $L \subset R$, this last tensor product is irreducible so that $W(\sigma) \simeq W(\tau_0) \otimes_Q L$ leading to the conclusion $L = Q$. We have thus to prove.

Proposition. Let ρ be an irreducible representation of k_Q over Q. Then $\rho \otimes_Q R$ is irreducible.

Proof. The Lie algebra k_Q splits over $Q(i)$. It follows that over $Q(i)$ all representations over C have equivalents. In particular this means that an irreducible representation ρ over Q decomposes over C into at most two representations. If ρ remains irreducible over $Q(i)$ hence over C, there is nothing to prove. Assume that $\rho \otimes_Q Q(i) \simeq \rho_1 \oplus \rho_2$ over $Q(i)$. If ρ_1 and ρ_2 are inequivalent, then the commutant of ρ is an algebra which when tensored with $Q(i)$ is isomorphic to $Q(i) \times Q(i)$. It follows that the commutant of $\rho(k_Q)$ in End $W(\rho)$ ($W(\rho)$ = representation space for ρ) is $Q(i)$. Since $Q(i) \otimes R \simeq C$ is a field, it follows that in this case too ρ remains irreducible. We have thus to consider now only the case

$$\rho \otimes Q(i) \simeq \sigma \oplus \sigma$$

two copies of the same irreducible representation. Let Δ be a simple system of $Q(i)$-roots with respect to T fixed as in the beginning of § 2 and w_0 be the Weyl group element defined there. Let $S \subset T$ be the maximal torus fixed pointwise by w_0. Let $\wedge$ be the highest weight of σ and $W(\wedge) \subset W(\sigma)$ the eigen space corresponding to $\wedge$. $W(\wedge)$ is defined over $Q(i)$. Let σ be considered as a subrepresenta-

tion through the direct sum decomposition over $Q(i)$ and label the two factors by 1, 2. Then we can choose the components so that we have

$$\overline{W(\Lambda)_1} \subset W(\sigma)_2, W(\rho \otimes_Q Q(i)) = W(\sigma)_1 \oplus W(\sigma)_2 \text{ and } W(\Lambda)_1 \subset W(\sigma)_1$$

is the highest weight space: otherwise $W(\sigma)_1$ would be stable under conjugation so that it will be defined over Q contradicting the irreducibility of ρ over Q. Similarly $\overline{W(\Lambda)_2} \subset W(\sigma)_1$. Now since complex conjugation takes t to t^{-1} in the torus we have necessarily $\overline{W(\Lambda)_1} = W(\Lambda^{-1})_2$. Since Λ^{-1} is necessarily the least weight of σ again, we conclude that $w_0(\Lambda) = \Lambda^{-1}$. Consider now the representation μ of the group B generated by w_0 and T on $E = W(\Lambda)_1 + W(\Lambda)_2 + W(\Lambda^{-1})_1 + W(\Lambda)_2$. We have then for $\mu(w_0)$, $\mu(w_0)^2$ is the unique element of order 2 in the group $\mu(T/S)(Q)$. Now μ is a 4-dimensional real irreducible representation of B as is easily seen. Its commuting algebra is thus a division algebra of degree 2. The restriction of $\text{End}_g(\rho)$ to E is seen to be nontrivial division algebra; since dim $E = 4$, these commuting algebras must coincide. If D denotes this division algebra E is necessarily a 1-dimensional vector space and the algebra generated by $B(Q)$ is contained in the commutant H of D in $\text{End}_Q(E) \subset M_4(Q)$. The last algebra is evidently isomorphic to D (note degree $D = 2$ so that $D \simeq D^0$). We will show that D is the definite quaternion algebra generated by i, j, k with $i^2 = j^2 = k^2 = -1$ $ij = k$, etc. To see this let L be the subfield of H generated by $\mu(t_Q)$. L is isomorphic $Q(i)$ where we denote by i the square root of -1 in L. Next set $j = \mu(w_0)$. Now $j^2 = \mu(w_0)^2$; it equals either the unique element of order 2 in L, viz., -1 or $j^2 = 1$. If $j^2 = 1$, $Q[j]$ contains a zero divisor a contradiction to $j \in H$. Thus $j^2 = -1$. Finally set $k = ij$. Then (ij). $(ij) = ij^2(j^{-1}ij) = ij^2 i^{-1} = j^2 = -1$. Showing that the algebra generated by $\mu(t_Q)$ and $\mu(w_0)$ is isomorphic to the definite quaternion algebra. This implies that D is a definite quaternion algebra over Q. Hence $D \otimes_Q R$ remains a division algebra proving that $\rho \otimes_Q R$ is irreducible.

References

[1] Borel A [1] 1963 *Topology* **2** 111–122

[2] Klingen H [1] 1955 *Math. Ann.* **129**

[3] Ramanathan K G 1961 *Math. Ann.* **143** 293–332

[4] Millson J and Raghunathan M S 1981 *Geometric construction of cohomology for arithmetic groups, in Geometry and Analysis*, Papers dedicated to V K Patodi, (Bangalore: Indian Academy of Sciences), p. 103; *Proc. Indian Acad. Sci.* (*Math Sci.*) **90** 103

[5] Siegel C L [1] 1943 *Am. J. Math.* **65** 1–86

[6] van der Waerden B L [1] 1964 *Modern algebra* (*English translation*) (New York: Frederic Ungar) Vol. II

Proc. Indian Acad. Sci. (Math. Sci.), Vol. 91, Number 2, July 1982, pp. 139-154.

Scattering of impulsive elastic waves by a fluid cylinder

B K RAJHANS and K M AGRAWAL*
Department of Physics and Mathematics, Indian School of Mines, Dhanbad 826004, India
* Department of Mathematics, Government Polytechnic, Dhanbad 826 001, India

MS received 7 December 1981

Abstract. We consider the scattering of impulsive SV waves by a fluid circular cylinder. The cylinder is embedded in an unbounded isotropic homogeneous elastic medium and it is filled with some acoustic fluid. The line source, generating the incident pulse is situated outside the cylinder parallel to its axis. We investigate the problem by the method of dual integral transformation as developed by Friedlander. The resulting integrals are evaluated approximately to obtain the short time estimate of the motion near the wave-front in the illuminated region of the elastic medium. We also interpret the approximate solution in terms of geometrical optics.

Keywords. Elastic waves; scattering; dual integral transformation; geometrical optics.

1. Introduction

The scattering and diffraction of two-dimensional elastic waves with a cylindrical obstacle in an unbounded medium has been considered in recent years. Gilbert and Knopoff (1959) discussed the scattering of impulsive elastic waves by a rigid circular cylinder situated in a homogeneous isotropic elastic medium. Gilbert (1960) considered the scattering of impulsive elastic waves by a smooth convex cylinder and obtained the formal solution of the problem using the technique of dual integral transformation developed by Friedlander (1954). An approximate evaluation of the solution was then obtained corresponding to the short-time behaviour of the scattered field near the wave-front both in the shadow as well as in the illuminated zone. Jha (1974) used the same technique to investigate the problem of diffraction of compressional waves by a fluid cylinder in a homogeneous medium. Rajhans and Mishra (1980) also considered the diffraction of impulsive elastic waves by a fluid cylinder in a homogeneous medium using the above technique.

In this paper, we investigate and discuss the scattering of the impulsive SV pulses by a circular cylinder filled with inviscid fluid material. The obstacle is supposed to be situated in an unbounded homogeneous isotropic elastic medium and the incident pulse is generated by a line source situated in the surrounding

elastic medium at a finite distance parallel to the axis of the cylinder. It is well-known that line stone formations usually exhibit a shear-wave arrival stronger than the compressional waves. Therefore in such cases shear waves are more prominent than the compressional waves (White 1965). We suppose that the velocities of P and SV waves outside the cylinder are α and β respectively and that of P waves inside the cylinder is α_0. To be specific, we assume $\alpha > \alpha_0 > \beta$. This assumption of the velocity distribution corresponds to the actual velocity distribution of elastic waves inside the earth and to the location of the source in the mantle and the outer core as the obstacle (Bullen 1963). The present investigation therefore throws some light on the scattering of SV waves by the core of the earth. We also suppose the density of the medium outside the cylinder is ρ and that inside the cylinder is ρ' where $\rho > \rho'$.

2. Formulation of the problem

Let the axis of the cylinder be taken as the z-axis and let a co-ordinate system (r, θ) be located in the (x, y) plane with $\theta = 0$, $r = r_0\ (> a)$ corresponding to the location of the line source which is parallel to the axis of the cylinder. The equation of the cylinder is $r = a$.

We define the elastic velocity potentials Φ_0, Φ and ψ corresponding to the wave equations inside and outside the cylinder respectively. Since z-axis is taken along the axis of the cylinder, the state of the media is fully determined if the velocity potentials Φ_0, Φ and ψ are obtained as a function of r, θ and t. It is thus required to find out the velocity potentials Φ_0, Φ and ψ as a function of r, θ and t which satisfy the wave equations

$$\frac{1}{\beta^2}\frac{\partial^2 \psi}{\partial t^2} - \nabla^2 \psi = \frac{2\pi}{r}\,\delta(r - r_0)\,\delta(t)\,\delta(\theta), \quad (r \geqslant a), \tag{1}$$

$$\frac{1}{\alpha^2}\frac{\partial^2 \Phi}{\partial t^2} - \nabla^2 \Phi = 0, \quad (r \geqslant a), \tag{2}$$

$$\frac{1}{\alpha_0^2}\frac{\partial^2 \Phi_0}{\partial t^2} - \nabla^2 \Phi_0 = 0, \quad (r \leqslant a), \tag{3}$$

where

$$\nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2};$$

the initial conditions

$$\begin{aligned} &\psi = \partial\psi/\partial t = 0, \text{ when } t = 0 \text{ except at } r = r_0,\ \theta = 0,\\ &\Phi = \partial\Phi/\partial t = 0, \text{ when } t = 0,\\ &\Phi_0 = \partial\Phi_0/\partial t = 0, \text{ when } t = 0, \end{aligned} \tag{4}$$

and the boundary conditions

$$\begin{aligned} &[T_{r\theta}]_{r=a+0} = 0,\\ &[T_{rr}]_{r=a+0} = [T_{rr}]_{r=a-0},\\ &[u_r]_{r=a+0} = [u_r]_{r=a-0}, \end{aligned} \tag{5}$$

where

$$T_{r\theta} = \nu \left\{ \frac{1}{r}\frac{\partial u_r}{\partial \theta} + \frac{\partial u_\theta}{\partial r} - \frac{u_\theta}{r} \right\},$$

$$T_{rr} = \lambda \frac{1}{a^2}\frac{\partial^2 \Phi}{\partial t^2} + 2\nu \frac{\partial u_r}{\partial r},$$

$$u_r = \frac{\partial \Phi}{\partial r} + \frac{1}{r}\frac{\partial \psi}{\partial \theta},$$

$$u_\theta = \frac{1}{r}\frac{\partial \Phi}{\partial \theta} - \frac{\partial \psi}{\partial r},$$

Here λ and ν are Lame's parameters and the symbol δ stands for Dirac delta function.

3. The formal solution

Let us define the Laplace transform $\bar{\psi}(r,\theta,s)$ of $\psi(r,\theta,t)$ by

$$\bar{\psi}(r,\theta,s) = \int_0^\infty \psi(r,\theta,t)\exp(-st)\,dt, \tag{6}$$

where s is the transform variable. Again we denote the Fourier transform ψ^* of $\bar{\psi}$ by

$$\psi^*(r,\mu,s) = \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^{\infty} \bar{\psi}(r,\theta,s)\exp(-i\mu\theta)\,d\theta \tag{7}$$

Applying these transformations to (1), (2) and (3) we get

$$\frac{d^2\psi^*}{dr^2} + \frac{1}{r}\frac{d\psi^*}{dr} - \left(\frac{s^2}{\beta^2} + \frac{\mu^2}{r^2}\right)\psi^* = -\frac{(2\pi)^{1/2}}{r}\delta(r - r_0), \quad (r \geqslant a) \tag{8}$$

$$\frac{d^2\Phi^*}{dr^2} + \frac{1}{r}\frac{d\Phi^*}{dr} - \left(\frac{s^2}{a^2} + \frac{\mu^2}{r^2}\right)\Phi^* = 0, \quad (r \geqslant a); \tag{9}$$

and

$$\frac{d^2\Phi_0^*}{dr^2} + \frac{1}{r}\frac{d\Phi_0^*}{dr} - \left(\frac{s^2}{a_0^2} + \frac{\mu^2}{r^2}\right)\Phi_0^* = 0, \ (r \leqslant a). \tag{10}$$

Also the Laplace–Fourier transformed boundary conditions are given by

$$[T_{r\theta}^*]_{r=a+0} = 0,$$

$$[T_{rr}^*]_{r=a+0} = [T_{rr}^*]_{r=a-0},$$

and

$$[u_r^*]_{r=a+0} = [u_r^*]_{r=a-0}. \tag{11}$$

It may be assumed that ψ^* is continuous at $r = r_0$.

Then (8) is equivalent to

$$\frac{d^2\psi^*}{dr^2} + \frac{1}{r}\frac{d\psi^*}{dr} - \left(\frac{s^2}{\beta^2} + \frac{\mu^2}{r^2}\right)\psi^* = 0; \quad (r \geqslant a), \tag{12}$$

and

$$[\psi^*]_{r_0-0}^{r_0+0} = 0, \; [d\psi^*/dr]_{r_0-0}^{r_0+0} = -(2\pi)^{1/2}/r_0. \tag{13}$$

After some mathematical calculation and Fourier inversion, we find that the Laplace transforms of the solution are given by

$$\bar{\psi}(r, \theta, s) = \int_{-\infty}^{\infty} I_{|\mu|}\left(\frac{sr}{\beta}\right) K_{|\mu|}\left(\frac{sr_0}{\beta}\right) \exp(i\mu\theta)\, d\mu$$

$$+ \int_{-\infty}^{\infty} K_\mu\left(\frac{sr}{\beta}\right) K_\mu\left(\frac{sr_0}{\beta}\right) \frac{L}{M} \exp(i\mu\theta)\, d\mu, \quad (r_0 \geqslant r \geqslant a) \tag{14}$$

$$\bar{\Phi}(r, \theta, s) = \int_{-\infty}^{\infty} K_\mu\left(\frac{sr}{\alpha}\right) K_\mu\left(\frac{sr_0}{\beta}\right) \frac{N}{M} \exp(i\mu\theta)\, d\mu \quad (r \geqslant a) \tag{15}$$

$$\bar{\Phi}_0(r, \theta, s) = \int_{-\infty}^{\infty} \frac{2\rho\mu^2 \beta^2/a^3 \cdot L + N \cdot P}{M \cdot Q} I_{|\mu|}\left(\frac{sr}{\alpha_0}\right)$$

$$\times K_\mu\left(\frac{sr_0}{\beta}\right) \exp(i\mu\theta)\, d\mu, \quad (r \leqslant a) \tag{16}$$

where

$$L = \frac{2\rho s^4}{a\beta\alpha_0} I'_\mu(sa/\alpha_0) I'_\mu(sa/\beta) K_\mu(sa/\alpha) - \frac{4\rho s^3 \beta}{a^2 \alpha\alpha_0} (I'_\mu)(sa/\alpha_0)$$

$$\times I'_\mu(sa/\beta) K'_\mu(sa/\alpha) - \frac{2\rho' s^4}{a\,\alpha\beta} I_\mu(sa/\alpha_0) I'_\mu(sa/\beta) K'_\mu(sa/\alpha)$$

$$- \frac{\rho s\beta^2}{\alpha_0}\left(\frac{s^2}{\beta^2} + \frac{2\mu^2}{a^2}\right) I'_\mu(sa/\alpha_0) I_\mu(sa/\beta) K_\mu(sa/\alpha)$$

$$+ \frac{2\rho s^4}{a\alpha\alpha_0} I'_\mu(sa/\alpha_0) I_\mu(sa/\beta) K'_\mu(sa/\alpha) + \frac{\rho' s^5}{\alpha\beta^2} I_\mu(sa/\alpha_0)$$

$$\times I_\mu(sa/\beta) K'_\mu(sa/\alpha) + \frac{4\rho\mu^2 s^3 \beta}{a^2 \alpha\alpha_0} I'_\mu(s_a/\alpha_0) I'_\mu(sa/\beta) K'_\mu(sa/\alpha)$$

$$+ \frac{4\rho s\mu^2 \beta^2}{a^4 \alpha_0} I'_\mu(sa/\alpha_0) I_\mu(sa/\beta) K_\mu(sa/\alpha)$$

$$+ \frac{2\rho' \mu^2 s^2}{a^3} I_\mu(sa/\alpha_0) I_\mu(sa/\beta) K_\mu(sa/\alpha),$$

$$M = \frac{4\rho s^3 \beta}{a^2 \alpha\alpha_0} I'_\mu(sa/\alpha_0) K'_\mu(sa/\alpha) K'_\mu(sa/\beta) - \frac{4\rho \beta\mu^2 s^3}{a^2 \alpha\alpha_0}$$

$$\times I'_\mu(sa/\alpha_0) K'_\mu(sa/\alpha) K'_\mu(sa/\beta) - \frac{4\rho s\mu^2 \beta^2}{a^4 \alpha_0} I'_\mu(sa/\alpha_0)$$

$$\times K_\mu(sa/\alpha) K_\mu(sa/\beta) - \frac{2\rho' \mu^2 s^2}{a^3} I_\mu(sa/\alpha_0) K_\mu(sa/\alpha) K_\mu(sa/\beta)$$

$$- \frac{2\rho s^4}{a\beta\alpha_0} I'_\mu(sa/\alpha_0) K_\mu(sa/\alpha) K'_\mu(sa/\beta) + \frac{2\rho' s^4}{a\,\alpha\beta} I_\mu(sa/\alpha_0)$$

$$\times K'_\mu(sa/\alpha) K'_\mu(sa/\beta) + \frac{\rho s \beta^2}{\alpha_0}\left(\frac{s^2}{\beta^2} + \frac{2\mu^2}{a^2}\right)^2 I'_\mu(sa/\alpha_0)$$

$$\times K_\mu(sa/\alpha) K_\mu(sa/\beta) - \frac{2\rho s^4}{a\,\alpha\alpha_0} I'_\mu(sa/\alpha_0) K'_\mu(sa/\alpha) K_\mu(sa/\beta)$$

$$- \frac{\rho' s^5}{\alpha\beta^2} I_\mu(sa/\alpha_0) K'_\mu(sa/\alpha) K_\mu\, sa/\beta),$$

$$N = \frac{4i\mu\rho s\beta^2}{a^4 \alpha_0} I'_\mu(sa/\alpha_0) + \frac{2i\mu\rho' s^2}{a^3} I_\mu(sa/\alpha_0)$$

$$- \frac{2i\mu\rho s\beta^2}{a^4\alpha_0}\left(\frac{s^2}{\beta^2} + \frac{2\mu^2}{a^2}\right) I'_\mu(sa/\alpha_0),$$

$$P = \frac{i\mu\rho\beta^2}{a}\left(\frac{s^2}{\beta^2} + \frac{2\mu^2}{a^2}\right) I_\mu\left(\frac{sa}{\beta}\right) K_\mu\left(\frac{sa}{\alpha}\right)$$

$$- \frac{2i\mu\rho\beta s^2}{a\alpha} I'_\mu\left(\frac{sa}{\beta}\right) K'_\mu\left(\frac{sa}{\alpha}\right)),$$

and

$$Q = \frac{i\mu\rho' s^2}{a} I_\mu\left(\frac{sa}{\beta}\right) I_\mu\left(\frac{sa}{\alpha_0}\right) + \frac{2i\mu\rho s\beta^2}{a^2\alpha_0} I_\mu\left(\frac{sa}{\beta}\right)$$

$$\times I'_\mu\left(\frac{sa}{\alpha_0}\right) - \frac{2i\mu\rho s^2\beta}{a\alpha_0} I'_\mu\left(\frac{sa}{\beta}\right) I'_\mu\left(\frac{sa}{\alpha_0}\right). \tag{17}$$

It is evident that (14), (15) and (16) give the integral representation of Laplace transform of the formal solution. The time solution can be obtained on performing Laplace inversion. But it is difficult to evaluate the integrals (14), (15) and (16) as they stand. However, if one is interested in short-time behaviour of the pulses, these integrals can be evaluated approximately for large, positive and real s.

4. Incident, reflected and refracted pulses

We first give a brief description of the geometry of the problem. Initially the incident sv pulse striking the outer surface of the cylinder gives rise to reflected S, reflected P and refracted P pulses according to the laws of ordinary geometrical optics (figure 1). When the rays strike the outer surface of the cylinder at the critical angle, the reflected P rays become tangential to the surface and as such they move along the surface. These surface waves at each point of their path give rise to lateral sv waves at critical angle in the outer medium and P waves along the tangent to the surface (Keller 1958). The former are denoted by $S(P)S$ and the latter by $S(P)P$ (Gilbert and Knopoff 1959). In the case of grazing incidence, the disturbances move along the surface and at each point of their path they shed diffracted sv waves in the outer medium tangential to the surface. These are denoted by $S(S)S$ (Bullen 1963).

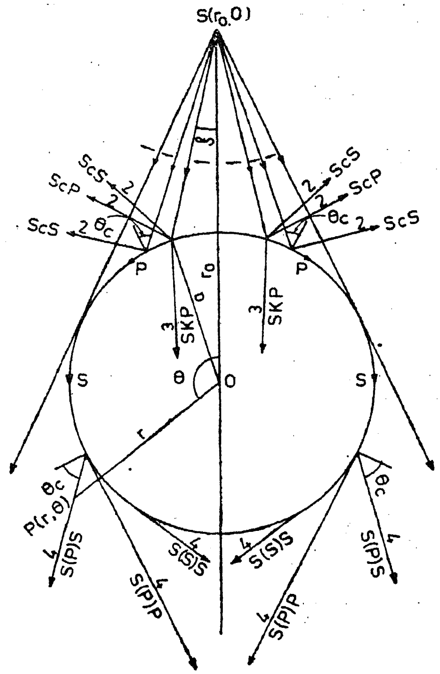


Figure 1. Incident sv pulse striking the outer surface of the cylinder. 1. Incident ray, 2. Reflected ray, 3. Refracted ray and 4. Diffracted ray.

Now we proceed to obtain the solution in the illuminated region of the elastic medium. For this purpose, we use the method of saddle point integration (Jeffreys and Jeffreys 1956). Therefore, we use the asymptotic approximations for modified Bessel functions occurring in the integrals (14), (15) and (16). We assume s to be large, real and positive. We know that for such s, the principal contributions to the integrals in (14), (15) and (16) arise from large values of $|\mu|$. Therefore using the various approximations for modified Bessel functions as given by (Mishra 1964a, b) in (14), (15) and (16), we find that

$$\bar{\psi}(r,\theta,s) \sim \int_{-\infty}^{\infty} f_1(\mu)\exp\{g_1(\mu)\}\,d\mu + \int_{-\infty}^{\infty} f_2(\mu)\exp\{g_2(\mu)\}\,d\mu, \quad (r_0 \geqslant r \geqslant a) \tag{18}$$

$$\bar{\Phi}(r,\theta,s) \sim \int_{-\infty}^{\infty} f_3(\mu)\exp\{g_3(\mu)\}\,d\mu, \quad (r \geqslant a) \tag{19}$$

$$\bar{\Phi}_0(r,\theta,s) \sim \int_{-\infty}^{\infty} f_4(\mu)\exp\{g_4(\mu)\}\,d\mu, \quad (r \leqslant a) \tag{20}$$

where

$$f_1(\mu) = \frac{\beta}{2}(\mu^2\beta^2 + s^2r^2)^{-1/4}(\mu^2\beta^2 + s^2r_0^2)^{-1/4},$$

$$g_1(\mu) = i\mu\theta + \left(\mu^2 + \frac{s^2r^2}{\beta^2}\right)^{1/2} - \mu\sinh^{-1}\left(\frac{\mu\beta}{sr}\right) - \left(\mu^2 + \frac{s^2r_0^2}{\beta^2}\right)^{1/2}$$
$$+ \mu\sinh^{-1}\left(\frac{\mu\beta}{sr_0}\right),$$

$$f_2(\mu) = \frac{\beta}{2}(\mu^2\beta^2 + s^2 r^2)^{-1/4}(\mu^2\beta^2 + s^2 r_0^2)^{-1/4}$$

$$\times \Big[4\rho\mu^2\beta^4\left(1+\frac{s^2a^2}{\mu^2\alpha^2}\right)^{1/2}\left(1+\frac{s^2a^2}{\mu^2\alpha_0^2}\right)^{1/2}\left(1+\frac{s^2a^2}{\mu^2\beta^2}\right)^{1/2}$$

$$-4\rho\mu^4\beta^4\left(1+\frac{s^2a^2}{\mu^2\alpha^2}\right)^{1/2}\left(1+\frac{s^2a^2}{\mu^2\alpha_0^2}\right)^{1/2}\left(1+\frac{s^2a^2}{\mu^2\beta^2}\right)^{1/2}$$

$$+2\mu\rho s^2a^2\beta^2\left(1+\frac{s^2a^2}{\mu^2\alpha_0^2}\right)^{1/2}\left(1+\frac{s^2a^2}{\mu^2\beta^2}\right)^{1/2}$$

$$+2\mu\rho' s^2a^2\beta^2\left(1+\frac{s^2a^2}{\mu^2\alpha^2}\right)^{1/2}\left(1+\frac{s^2a^2}{\mu^2\beta^2}\right)^{1/2}$$

$$-\rho\beta^4a^4\left(\frac{s^2}{\beta^2}+\frac{2\mu^2}{a^2}\right)^2\left(1+\frac{s^2a^2}{\mu^2\alpha_0^2}\right)^{1/2}$$

$$-2\mu\rho s^2a^2\beta^2\left(1+\frac{s^2a^2}{\mu^2\alpha^2}\right)^{1/2}\left(1+\frac{s^2a^2}{\mu^2\alpha_0^2}\right)^{1/2}$$

$$+4\rho\mu^2\beta^4\left(1+\frac{s^2a^2}{\mu^2\alpha_0^2}\right)^{1/2} - \rho' s^4a^4\left(1+\frac{s^2a^2}{\mu^2\alpha^2}\right)^{1/2}$$

$$+2\mu\rho' s^2a^2\beta^2\Big]$$

$$\Big[4\rho\mu^2\beta^4\left(1+\frac{s^2a^2}{\mu^2\alpha^2}\right)^{1/2}\left(1+\frac{s^2a^2}{\mu^2\alpha_0^2}\right)^{1/2}\left(1+\frac{s^2a^2}{\mu^2\beta^2}\right)^{1/2}$$

$$-4\rho\mu^4\beta^4\left(1+\frac{s^2a^2}{\mu^2\alpha^2}\right)^{1/2}\left(1+\frac{s^2a^2}{\mu^2\alpha_0^2}\right)^{1/2}\left(1+\frac{s^2a^2}{\mu^2\beta^2}\right)^{1/2}$$

$$+2\mu\rho s^2a^2\beta^2\left(1+\frac{s^2a^2}{\mu^2\alpha_0^2}\right)^{1/2}\left(1+\frac{s^2a^2}{\mu^2\beta^2}\right)^{1/2}$$

$$+2\mu\rho' s^2a^2\beta^2\left(1+\frac{s^2a^2}{\mu^2\alpha^2}\right)^{1/2}\left(1+\frac{s^2a^2}{\mu^2\beta^2}\right)^{1/2}$$

$$+\rho\beta^4a^4\left(\frac{s^2}{\beta^2}+\frac{2\mu^2}{a^2}\right)^2\left(1+\frac{s^2a^2}{\mu^2\alpha_0^2}\right)^{1/2}$$

$$+2\mu\rho s^2a^2\beta^2\left(1+\frac{s^2a^2}{\mu^2\alpha^2}\right)^{1/2}\left(1+\frac{s^2a^2}{\mu^2\alpha_0^2}\right)^{1/2}$$

$$-4\rho\mu^2\beta^4\left(1+\frac{s^2a^2}{\mu^2\alpha_0^2}\right)^{1/2} + \rho' s^4a^4\left(1+\frac{s^2a^2}{\mu^2\alpha^2}\right)^{1/2}$$

$$-2\mu\rho' s^2a^2\beta^2\Big]$$

$$g_2(\mu) = i\mu\theta + 2\left(\mu^2 + \frac{s^2 a^2}{\beta^2}\right)^{1/2} - 2\mu \sinh^{-1}\left(\frac{\mu\beta}{sa}\right)$$
$$- \left(\mu^2 + \frac{s^2 r^2}{\beta^2}\right)^{1/2} + \mu \sinh^{-1}\left(\frac{\mu\beta}{sr}\right) - \left(\mu^2 + \frac{s^2 r_0^2}{\beta^2}\right)^{1/2}$$
$$+ \mu \sinh^{-1}\left(\frac{\mu\beta}{sr_0}\right),$$

$$f_3(\mu) = \frac{2i\beta^2\left(1 + \frac{s^2 a^2}{\mu^2 \alpha^2}\right)^{1/4}\left(1 + \frac{s^2 a^2}{\mu^2 \beta^2}\right)^{1/4}}{\left(1 + \frac{s^2 r^2}{\mu^2 \alpha^2}\right)^{1/4}\left(1 + \frac{s^2 r_0^2}{\mu^2 \beta^2}\right)^{1/4}}$$
$$\times \left\{(2\mu\rho\beta^2\left(1 + \frac{s^2 a^2}{\mu^2 \alpha_0^2}\right)^{1/2} + \rho' s^2 a^2 - \mu\rho\beta^2 a^2\left(\frac{s^2}{\beta^2} + \frac{2\mu^2}{a^2}\right)\right.$$
$$\left.\left(1 + \frac{s^2 a^2}{\mu^2 \alpha_0^2}\right)^{1/2}\right\}$$

$$\left\{4\rho\mu^2\beta^4\left(1 + \frac{s^2 a^2}{\mu^2 \alpha^2}\right)^{1/2}\left(1 + \frac{s^2 a^2}{\mu^2 \alpha_0^2}\right)^{1/2}\left(1 + \frac{s^2 a^2}{\mu^2 \beta^2}\right)^{1/2}\right.$$
$$- 4\rho\mu^4\beta^4\left(1 + \frac{s^2 a^2}{\mu^2 \alpha^2}\right)^{1/2}\left(1 + \frac{s^2 a^2}{\mu^2 \alpha_0^2}\right)^{1/2}\left(1 + \frac{s^2 a^2}{\mu^2 \beta^2}\right)^{1/2}$$
$$+ 2\mu\rho s^2 a^2 \beta^2\left(1 + \frac{s^2 a^2}{\mu^2 \alpha_0^2}\right)^{1/2}\left(1 + \frac{s^2 a^2}{\mu^2 \beta^2}\right)^{1/2}$$
$$+ 2\mu\rho' s^2 a^2 \beta^2\left(1 + \frac{s^2 a^2}{\mu^2 \alpha^2}\right)^{1/2}\left(1 + \frac{s^2 a^2}{\mu^2 \beta^2}\right)^{1/2}$$
$$+ \rho\beta^4 a^4\left(\frac{s^2}{\beta^2} + \frac{2\mu^2}{a^2}\right)^2\left(1 + \frac{s^2 a^2}{\mu^2 \alpha_0^2}\right)^{1/2}$$
$$+ 2\mu\rho s^2 a^2 \beta^2\left(1 + \frac{s^2 a^2}{\mu^2 \alpha^2}\right)^{1/2}\left(1 + \frac{s^2 a^2}{\mu^2 \alpha_0^2}\right)^{1/2}$$
$$- 4\rho\mu^2\beta^4\left(1 + \frac{s^2 a^2}{\mu^2 \alpha_0^2}\right)^{1/2} + \rho' s^4 a^4\left(1 + \frac{s^2 a^2}{\mu^2 \alpha^2}\right)^{1/2}$$
$$\left.- 2\mu\rho' s^2 a^2 \beta^2\right\},$$

$$g_3(\mu) = i\mu\theta + \left(\mu^2 + \frac{s^2 a^2}{\alpha^2}\right)^{1/2} - \mu \sinh^{-1}\left(\frac{\mu\alpha}{sa}\right)$$
$$+ \left(\mu^2 + \frac{s^2 a^2}{\beta^2}\right)^{1/2} - \mu \sinh^{-1}\left(\frac{\mu\beta}{sa}\right) - \left(\mu^2 + \frac{s^2 r^2}{\alpha^2}\right)^{1/2}$$
$$+ \mu \sinh^{-1}\left(\frac{\mu\alpha}{sr}\right) - \left(\mu^2 + \frac{s^2 r_0^2}{\beta^2}\right)^{1/2} + \mu \sinh^{-1}\left(\frac{\mu\beta}{sr_0}\right),$$

$$f_4(\mu) = -2i\beta\alpha_0 a^2 \frac{\left(1+\frac{s^2a^2}{\mu^2\beta^2}\right)^{1/4}\left(1+\frac{s^2a^2}{\mu^2\alpha_0^2}\right)^{1/4}}{\left(1+\frac{s^2r^2}{\mu^2\alpha_0^2}\right)^{1/4}\left(1+\frac{s^2r_0^2}{\mu^2\beta^2}\right)^{1/4}}$$

$$\times \{2\mu\alpha\rho^2\beta^2 s^2(\mu^2\beta^2+s^2a^2)^{1/2}(\mu^2\alpha_0^2+s^2a^2)^{1/2}$$

$$-2\mu\rho^2 s^2\beta^3(\mu^2\alpha^2+s^2a^2)^{1/2}(\mu^2\alpha_0^2+s^2a^2)^{1/2}$$

$$-\rho\rho'\mu\beta\alpha_0 a^2 s^4(\mu^2\alpha^2+s^2a^2)^{1/2} - 2\mu\alpha\rho^2 s^2\beta^3$$

$$\times(\mu^2\alpha_0^2+s^2a^2)^{1/2} - \rho\rho'\mu\alpha\alpha_0\beta a^2 s^4 + 2\mu\rho^2\beta^2 s^2$$

$$\times(\mu^2\alpha^2+s^2a^2)^{1/2}(\mu^2\beta^2+s^2a^2)^{1/2}(\mu^2\alpha_0^2+s^2a^2)^{1/2}\}$$

$$\{4\rho\beta^3(\mu^2\alpha^2+s^2a^2)^{1/2}(\mu^2\beta^2+s^2a^2)^{1/2}(\mu^2a_0^2+s^2a^2)^{1/2}$$

$$-4\rho\mu^2\beta^3(\mu^2\alpha^2+s^2a^2)^{1/2}(\mu^2\beta^2+s^2a^2)^{1/2}(\mu^2\alpha_0^2+s^2a^2)^{1/2}$$

$$-4\rho\alpha\mu^2\beta^4(\mu^2\alpha_0^2+s^2a^2)^{1/2} - 2\rho'\alpha\alpha_0\mu^2 s^2a^2\beta^2$$

$$+2\rho\alpha\beta s^2a^2(\mu^2\beta^2+s^2a^2)^{1/2}(\mu^2\alpha_0^2+s^2a^2)^{1/2}$$

$$+2\rho'\alpha_0\beta s^2a^2(\mu^2\alpha^2+s^2a^2)^{1/2}(\mu^2\beta^2+s^2a^2)^{1/2}$$

$$+\rho\alpha a^4\beta^4\left(\frac{s^2}{\beta^2}+\frac{2\mu^2}{a^2}\right)^2(\mu^2\alpha_0^2+s\,a^2)^{1/2}$$

$$+\rho'\alpha_0 s^4a^4(\mu^2\alpha^2+s^2a^2)^{1/2} + 2\rho\, s^2a^2\beta^2(\mu^2\alpha^2+s^2a^2)^{1/2}$$

$$\{(\mu^2\alpha_0^2+s^2a^2)^{1/2}\}\times\{\rho'\alpha_0 s^2a^2 + 2\rho\beta^2(\mu^2\alpha_0^2+s^2a^2)^{1/2}$$

$$-2\rho\beta(\mu^2\beta^2+s^2a^2)^{1/2}(\mu^2\alpha_0^2+s^2a^2)^{1/2}\},$$

and

$$g_4(\mu) = i\mu\theta + \left(\mu^2+\frac{s^2a^2}{\beta^2}\right)^{1/2} - \mu\sinh^{-1}\left(\frac{\mu\beta}{sa}\right) - \left(\mu^2+\frac{s^2a^2}{\alpha_0^2}\right)^{1/2}$$

$$+\mu\sinh^{-1}\left(\frac{\mu\alpha_0}{sa}\right) + \left(\mu^2+\frac{s^2r^2}{\alpha_0^2}\right)^{1/2} - \mu\sinh^{-1}\left(\frac{\mu\alpha_0}{sr}\right)$$

$$-\left(\mu^2+\frac{s^2r_0^2}{\beta^2}\right)^{1/2} + \mu\sinh^{-1}\left(\frac{\mu\beta}{sr_0}\right). \tag{21}$$

Evaluating the integrals in (18), (19) and (20), we find that the saddle points of all the four integrals are at $i\lambda s$ where

$$\lambda = \frac{r_0\sin\xi}{\beta}.$$

The first integral in (18) has a saddle point μ_0 where μ_0 is the solution of

$$i\theta + \sinh^{-1}\left(\frac{\mu_0\beta}{sr_0}\right) - \sinh^{-1}\left(\frac{\mu_0\beta}{sr}\right) = 0 \tag{22}$$

We can solve this equation geometrically. Let S be the source and $P_1(r, \theta)$ a point in the illuminated region (figure 2)

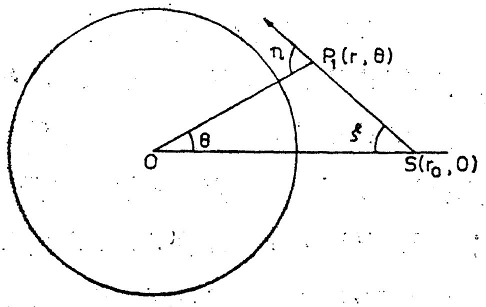


Figure 2. Geometrical interpretation of the saddle point for the incident pulse.

Let

$$\angle OSP_1 = \zeta \text{ and } \pi - \angle OP_1S = \eta$$

then

$$\theta = \eta - \xi, \ \sin\ \eta = \frac{r_0 \sin \xi}{r} \tag{23}$$

Now $\sinh^{-1}\left(\frac{\mu_0 \beta}{s r_0}\right) = i\xi$ defines the solution of (22) because from (23) we have

$$\sinh^{-1}\left(\frac{\mu_0 \beta}{s r}\right) = i\eta \tag{24}$$

and therefore (22) reduces to the first equation of (23). Thus we see that the saddle point μ_0 of the first integral in (18) can be associated with the incident ray.

At the saddle point μ_0 where

$$\mu_0 = i\frac{r_0 \sin \xi}{\beta} s,$$

we have

$$f_1(\mu_0) = \beta / \{2(r_0 \cos \xi)^{1/2} (r^2 - r_0^2 \sin^2 \xi)^{1/4} s$$

$$g_1(\mu_0) = -\frac{s}{\beta} R_1.$$

$$g_1''(\mu_0) = -\beta R_1 / \{r_0 \cos \xi (r^2 - r_0^2 \sin^2 \xi)^{1/2} s\} \tag{25}$$

where $R_1 = r_0 \cos \xi - r \cos \eta$ is the distance between the source and the receiver on the incident ray path. We find that the line of steepest descent through the saddle point is parallel to the real axis in the μ-plane since $g_1''(\mu_0)$ is negative. Therefore performing saddle point integration and using (25) we obtain (Jeffreys and Jeffreys 1956)

$$\int_{-\infty}^{\infty} f_1(\mu) \exp\{g_1(\mu)\} d\mu \sim \left(\frac{\pi\beta}{2R_1 s}\right)^{1/2} \exp -(sR_1)/\beta. \tag{26}$$

The saddle point μ_0 of the second integral in (18) is the solution of

$$i\theta - 2\sinh^{-1}\left(\frac{\mu_0\beta}{sa}\right) + \sinh^{-1}\left(\frac{\mu_0\beta}{sr_0}\right) + \sinh^{-1}\left(\frac{\mu_0\beta}{sr}\right) = 0. \tag{27}$$

We will interpret this geometrically. Let $S(r_0, 0)$ be the source and $P_2(r, \theta)$ be a point in the illuminated zone (figure 3).

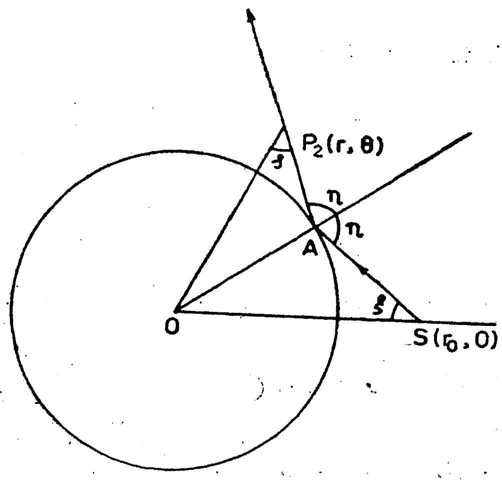


Figure 3. Geometrical interpretation of the saddle point for the reflected 'S' pulse.

Let $\angle OSA = \xi$, $\pi - \angle OAP_2 = \eta = \pi - \angle OAS$
and $\angle OP_2A = \zeta$ (28)
then

$$\theta = 2\eta - \xi - \zeta \sin\eta = \frac{r_0 \sin\xi}{a},\ \sin\zeta = \frac{a}{r}\sin\eta \tag{29}$$

Hence we can solve (27) by setting

$$\sinh^{-1}\frac{\mu_0\beta}{sr_0} = i\xi,\ \sinh^{-1}\frac{\mu_0\beta}{sa} = i\eta,\ \sinh^{-1}\frac{\mu_0\beta}{sr} = i\zeta. \tag{30}$$

Therefore the saddle point μ_0 of the second integral in (18) can be associated with the reflected S ray. At the saddle point we have

$$f_2(\mu_0) = \frac{\beta}{2s(rr_0\cos\xi\cos\zeta)^{1/2}} \times \frac{\{4\rho\beta\sin^2\eta\cos\eta(1-n^2\sin^2\eta)^{1/2}(1-n_1^2\sin^2\eta)^{1/2} - \rho\alpha(1-2\sin^2\eta)^2(1-n_1^2\sin^2\eta)^{1/2} - \rho'\alpha_0(1-n^2\sin^2\eta)^{1/2}\}}{\{4\rho\beta\sin^2\eta\cos\eta(1-n^2\sin^2\eta)^{1/2}(1-\eta_1^2\sin^2\eta)^{1/2} + \rho\alpha(1-2\sin^2\eta)^2(1-n_1{}^2\sin^2\eta)^{1/2} + \rho'\alpha_0(1-n^2\sin^2\eta)^{1/2}\}},$$

$$g_2(\mu_0) = -\left(\frac{R_2+R_3}{\beta}\right)s,$$

$$g_2''(\mu_0) = -\frac{\beta(rR_2\cos\zeta + r_0R_3\cos\xi)}{sarr_0\cos\xi\cos\eta\cos\zeta} \tag{31}$$

where $n = a/\beta$,

$n_1 = a_0/\beta$,

$R_2 = r_0 \cos \xi - a \cos \eta = SA$

$R_3 = r \cos \zeta - a \cos \eta = AP_2$.

Performing saddle point integration and using (31) we obtain

$$\int_{-\infty}^{\infty} f_2(\mu) \exp\{g_2(\mu)\}\, d\mu \sim \frac{\{4\rho\beta \sin^2\eta \cos\eta\,(1-n^2\sin^2\eta)^{1/2}(1-n_1^2\sin^2\eta)^{1/2} - \rho a(1-2\sin^2\eta)^2(1-n_1^2\sin^2\eta)^{1/2} - \rho' a_0(1-n^2\sin^2\eta)^{1/2}\}}{\{4\rho\beta \sin^2\eta \cos\eta\,(1-n^2\sin^2\eta)^{1/2}(1-n_1^2\sin^2\eta)^{1/2} + \rho a(1-2\sin^2\eta)^2(1-n_1^2\sin^2\eta)^{1/2} + \rho' a_0(1-n^2\sin^2\eta^{1/2}\}},$$

$$\times \left\{\frac{\pi a\beta \cos\eta}{2s(rR_2\cos\zeta + r_0R_3\cos\xi)}\right\}^{1/2} \exp -\left(\frac{R_2+R_3}{\beta}\right)s. \tag{32}$$

Now it remains to evaluate the integrals in (19) and (20). As before we see that the saddle point μ_0 of the integral in (19) is the solution of

$$i\theta + \sinh^{-1}\left(\frac{\mu_0 a}{sr}\right) + \sinh^{-1}\left(\frac{\mu_0 \beta}{sr_0}\right) - \sinh^{-1}\left(\frac{\mu_0 a}{sa}\right) - \sinh^{-1}\left(\frac{(\mu_0 \beta}{sa}\right) = 0. \tag{33}$$

This again can be solved geometrically (figure 4). Let $S(r_0, 0)$ be the source, SA the incident ray and AP_3 the reflected P ray. Let the coordinates of the point P_3 be (r, θ). Let

$$\angle \mathrm{OSA} = \xi, \pi - \angle \mathrm{OAS} = \eta, \pi - \angle \mathrm{OAP_3} = \delta', \quad \angle \mathrm{OP_3A} = \zeta' \tag{34}$$

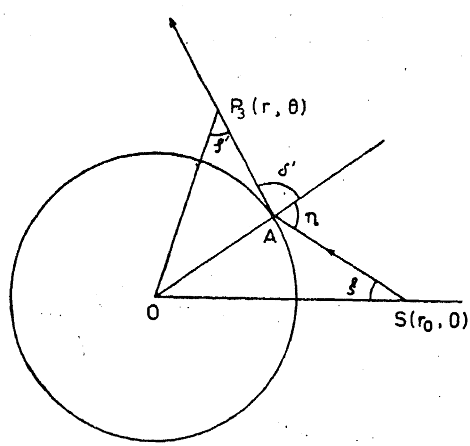


Figure 4. Geometrical interpretation of the saddle point for the feflected 'P' pulse.

then

$$\theta = \eta + \delta' - \xi - \zeta', \quad r_0 \sin \xi = a \ \sin \eta, r \sin \zeta' = a \sin \delta',$$

$$\sin \eta = \beta/\alpha \sin \delta'. \tag{35}$$

Thus the saddle point μ_0 of (19) is related to the reflected P pulse. **Hence** performing saddle point integrating as before one obtains

$$\int_{-\infty}^{\infty} f_3(\mu) \exp \{g_3(\mu)\} \, d\mu$$

$$\sim \{4\rho\alpha \sin \eta \cos \eta \, (1 - n_1^2 \sin^2 \eta)^{1/2} \ (1 - 2 \sin^2 \eta)\}$$

$$\{4\rho\beta \sin^2 \eta \cos \eta \, (1 - n^2 \sin^2 \eta)^{1/2} (1 - n_1^2 \sin^2 \eta)^{1/2}$$

$$+ \rho\alpha \, (1 - 2 \sin^2 \eta)^2 \, (1 - n_1^2 \sin^2 \eta)^{1/2} + \rho' \, \alpha_0 \, (1 - n^2 \sin^2 \eta)^{1/2}\}$$

$$\times \left\{\frac{\pi a \, \beta^2 \cos^2 \delta'}{2s \, (\beta R_2 \, r \cos \delta' \cos \zeta' + \alpha \, R_4 r_0 \cos \xi \cos \eta)}\right\}^{1/2}$$

$$\times \exp - \left(\frac{R_2}{\beta} + \frac{R_4}{\alpha}\right) s \tag{36}$$

where

$$R_3 = r_0 \cos \xi - a \cos \eta$$

$$R_4 = r \cos \zeta' - a \cos \delta'$$

The saddle point μ_0 of the integral in (20) is the solution of

$$i\theta + \sinh^{-1} \left(\frac{\mu_0 \beta}{s r_0}\right) + \sinh^{-1} \left(\frac{\mu_0 \alpha_0}{s a}\right)$$

$$- \sinh^{-1} \left(\frac{\mu_0 \alpha_0}{s r}\right) - \sinh^{-1} \left(\frac{\mu_0 \beta}{s a}\right) = 0. \tag{37}$$

Let $S(r_0, 0)$ be the source, SA the incident ray and AP_4 the refracted P ray Let the coordinates of P_4 be (r, θ) (figure 5). Let

$$\angle OSA = \xi, \ \pi - \angle OAS = \eta, \ \pi - \angle OP_4A = \zeta'' \ \angle OAP_4 = \delta'' \tag{38}$$

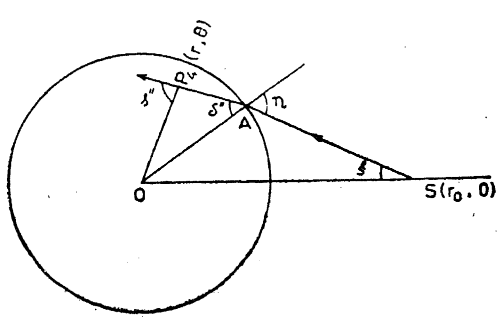


Figure 5. Geometrical interpretation of the saddle point for the refracted 'P' pulse

then

$$\theta = \eta + \zeta'' - \xi - \delta'',\ \sin\eta = \frac{r_0 \sin\xi}{a},\ \sin\delta'' = \alpha_0/\beta \sin\eta,$$

$$\sin\zeta'' = a/r \sin\delta''. \tag{39}$$

Therefore (37) can be solved by putting

$$\sinh^{-1}\left(\frac{\mu_0 \beta}{s r_0}\right) = i\xi,\ \sinh^{-1}\left(\frac{\mu_0 \alpha_0}{sa}\right) = i\delta''$$

$$\sinh^{-1}\left(\frac{\mu_0 \alpha_0}{sr}\right) = i\zeta'',\ \sinh^{-1}\left(\frac{\mu_0 \beta}{sa}\right) = i\eta. \tag{40}$$

Hence the saddle point μ_0 of (20) is related to the refracted P pulse. Therefore if we approximate (20) by the method of steepest descents, we obtain

$$\int_{-\infty}^{\infty} f_4(\mu) \exp\{g_4(\mu)\}\, d\mu \sim$$

$$-\frac{\{4\rho\alpha_0 \sin\eta \cos\eta\,(1 - n^2\sin^2\eta)^{1/2}\}}{\{4\rho\beta \sin^2\eta \cos\eta\,(1 - n^2\sin^2\eta)^{1/2}(1 - n_1^2\sin^2\eta)^{1/2} + \rho\alpha(1 - 2\sin^2\eta)^2(1 - n_1^2\sin^2\eta)^{1/2} + \rho'\alpha_0(1 - n^2\sin^2\eta)^{1/2}\}}$$

$$\times \left\{\frac{\pi\alpha\beta^2\cos^2\delta''}{2s(\beta R_2 r\cos\zeta''\cos\delta'' + \alpha_0 R_5 r_0 \cos\xi\cos\eta)}\right\}^{1/2}$$

$$\times \exp -\left(\frac{R_2}{\beta} + \frac{R_5}{\alpha_0}\right)s \tag{41}$$

where

$$R_5 = a\cos\delta'' - r\cos\zeta'' = AP_4.$$

Substituting the results (26) and (32) in (18) we get

$$\bar{\psi}(r,\theta;s) \sim \left(\frac{\pi\beta}{2R_1 s}\right)^{1/2} \exp -(st_1)$$

$$+ \frac{\{4\rho\beta\sin^2\eta\cos\eta\,(1 - n^2\sin^2\eta)^{1/2}(1 - n_1^2\sin^2\eta)^{1/2} - \rho\alpha(1 - 2\sin^2\eta)^2(1 - n_1^2\sin^2\eta)^{1/2} - \rho'\alpha_0(1 - n^2\sin^2\eta)^{1/2}\}}{\{4\rho\beta\sin^2\eta\cos\eta\,(1 - n^2\sin^2\eta)^{1/2}(1 - n_1^2\sin^2\eta)^{1/2} + \rho\alpha(1 - 2\sin^2\eta)^2(1 - n_1^2\sin^2\eta)^{1/2} + \rho'\alpha_0(1 - n^2\sin^2\eta)^{1/2}\}}$$

$$\times \left\{\frac{\pi\alpha\beta\cos\eta}{2s(r R_2\cos\zeta + r_0 R_3\cos\xi)}\right\}^{1/2} \exp -(st_2),$$

$$(r_0 \geqslant r \geqslant a) \tag{42}$$

where t_1 and t_2 are respectively the arrival times of the incident and reflected S pulses at the point (r, θ).

If we use (36) and (41) in (19) and (20) respectively, we get

$$\bar{\Phi}(r,\theta,s) \sim \frac{4\rho\alpha \sin\eta\cos\eta\,(1 - n_1^2\sin^2\eta)^{1/2}(1 - 2\sin^2\eta)}{\{4\rho\beta\sin^2\eta\cos\eta\,(1 - n^2\sin^2\eta)^{1/2}(1 - n_1^2\sin^2\eta)^{1/2} + \rho\alpha(1 - 2\sin^2\eta)^2(1 - n_1^2\sin^2\eta)^{1/2} + \rho'\alpha_0(1 - n^2\sin^2\eta)^{1/2}\}}$$

$$\times \left\{ \frac{\pi a \beta^2 \cos^2 \delta'}{2s\,(\beta R_2 r \cos \zeta' \cos \delta' + \alpha R_4 r_0 \cos \xi \cos \eta)} \right\}^{1/2}$$

$$\times \exp - (s t_3), \ (r \geqslant a) \tag{43}$$

$$\bar{\Phi}_0 (r, \theta, s) \sim - \frac{4 \rho \alpha_0 \sin \eta \cos \eta \,(1 - n^2 \sin^2 \eta)^{1/2}}{\{4 \rho \beta \sin^2 \eta \cos \eta \,(1 - n^2 \sin^2 \eta)^{1/2} (1 - n_1^2 \sin^2 \eta)^{1/2}}$$

$$+ \rho \alpha \,(1 - 2 \sin^2 \eta)^2 (1 - n_1{}^2 \sin^2 \eta)^{1/2} + \rho' \alpha_0 (1 - n^2 \sin^2 \eta)^{1/2}\}$$

$$\times \left\{ \frac{\pi a \beta^2 \cos^2 \delta''}{2s\,(\beta R_2 r \cos \zeta'' \cos \delta'' + \alpha_0 R_5 r_0 \cos \xi \cos \eta)} \right\}^{1/2}$$

$$\times \exp - (s t_4), \qquad (r \leqslant a) \tag{44}$$

where t_3 and t_4 are respectively the arrival times of the reflected and refracted P pulses at the point (r, θ).

Now, we can obtain the short-time approximations for the solution by performing Laplace inversion. We use the known result of Laplace inversion (Churchill (1958)

$$\frac{1}{2\pi i} \int_{\epsilon - i\infty}^{\epsilon + i\infty} s^{-\mu} \exp \{(t - T) s\}\, ds = \frac{(t - T)^{\mu - 1}}{\Gamma(\mu)} H(t - T) \tag{45}$$

where $\mu > 0$ and Γ stands for Gamma function.
Thus we have

$$\psi (r, \theta, t) \sim \frac{1}{\{2 t_1 (t - t_1)\}^{1/2}} H (t - t_1)$$

$$\{4 \rho \beta \sin^2 \eta \cos \eta \ (1 - n^2 \sin^2 \eta)^{1/2} \ (1 - n_1{}^2 \sin^2 \eta)^{1/2}$$

$$+ \frac{- \rho \alpha \,(1 - 2 \sin^2 \eta)^2 (1 - n_1^2 \sin^2 \eta)^{1/2} - \rho' \alpha_0 (1 - n^2 \sin^2 \eta)^{1/2}\}}{\{4 \rho \beta \sin^2 \eta \cos \eta \,(1 - n^2 \sin^2 \eta)^{1/2} (1 - n_1^2 \sin \eta)^{1/2}}$$

$$+ \rho \alpha \,(1 - 2 \sin^2 \eta)^2 (1 - n_1^2 \sin^2 \eta)^{1/2} + \rho' \alpha_0 (1 - n^2 \sin^2 \eta)^{1/2}\}$$

$$\times \left\{ \frac{a \beta \cos \eta}{2\,(r R_2 \cos \zeta + r_0 R_3 \cos \xi)} \right\}^{1/2} \frac{H (t - t_2)}{(t - t_2)^{1/2}}, \qquad (r_0 \geqslant r \geqslant a) \tag{46}$$

$$\Phi (r, \theta, t) \sim \frac{4 \rho \alpha \sin \eta \cos \eta \,(1 - n_1{}^2 \sin^2 \eta)^{1/2} \ (1 - 2 \sin^2 \eta)}{\{4 \rho \beta \sin^2 \eta \cos \eta \,(1 - n^2 \sin^2 \eta)^{1/2} (1 - n_1^2 \sin^2 \eta)^{1/2}}$$

$$+ \rho \alpha \,(1 - 2 \sin^2 \eta)^2 (1 - n_1^2 \sin^2 \eta)^{1/2} + \rho' \alpha_0 (1 - n^2 \sin^2 \eta)^{1/2}\}$$

$$\times \left\{ \frac{a \beta^2 \cos^2 \delta'}{2(\beta R_2 r \cos \zeta' \cos \delta' + \alpha R_4 r_0 \cos \xi \cos \eta)} \right\}^{1/2}$$

$$\times \frac{H (t - t_3)}{(t - t_3)^{1/2}}, \qquad (r \geqslant a) \tag{47}$$

$$\Phi_0 (r, \theta, t) \sim - \frac{4 \rho \alpha_0 \sin \eta \cos \eta \,(1 - n^2 \sin^2 \eta)^{1/2}}{\{4 \rho \beta \sin^2 \eta \cos \eta \,(1 - n^2 \sin^2 \eta)^{1/2} (1 - n_1^2 \sin^2 \eta)^{1/2}}$$

$$+ \rho \alpha \,(1 - 2 \sin^2 \eta)^2 (1 - n_1^2 \sin^2 \eta)^{1/2} + \rho' \alpha_0 (1 - n^2 \sin^2 \eta)^{1/2}\}$$

$$\times \left\{\frac{a\beta^2 \cos^2 \delta''}{2\,(\beta R_2 r \cos \zeta'' \cos \delta'' + a_0 R_5 r_0 \cos \xi \cos \eta)}\right\}^{1/2}$$

$$\times \frac{H(t-t_4)}{(t-t_4)^{1/2}}, \qquad (r \leqslant a) \tag{48}$$

We can physically interpret the solutions (46), (47) and (48) in the following way (Friedlander 1958). The first term in (46) represents the incident pulse. The second term in (46) represents the reflected sv pulse. Its first factor is the reflection coefficient for a plane boundary (Brekhovshikh 1960). The second and the third factors may be interpreted as the divergence and source factors respectively. The approximation for the reflected P pulse is given by (47). Its first factor is the reflection coefficient for a plane boundary (Brekhovshikh 1960). The second may be interpreted as the divergence factor due to the curved surface and the third as the source term. Similarly (48) represents the refracted P pulse. The first factor is the refraction coefficient for a plane boundary (Brekhovshikh 1960) and the second and third factors may be interpreted as the divergence and source terms respectively.

5. Conclusion

We have obtained the solution and interpreted them in terms of geometrical optics as the incident, reflected and refracted pulses. We note that the arrival time of the reflected pulses is the same whether the obstacle is a rigid, weak or fluid cylinder. However, we have not obtained the pulses associated with multiple reflection and refraction as it becomes very lengthy. We expect to discuss this case separately.

Acknowledgement

The authors are grateful to the authorities of the Indian School of Mines, Dhanbad, for the financial support and other facilities provided during the preparation of this paper.

References

[1] Brekhovskikh L M 1960 *Waves in layered media* (New York : Academic Press)
[2] Bullen K E 1963 *Theory of seismology* (Cambridge University Press)
[3] Churchill R V 1958 *Operational mathematics*, New York
[4] Friedlander F G 1954 *Commun. Pure Appl. Math.* **7** 705
[5] Friedlander F G 1958 *Sound pulses* (Cambridge University Press)
[6] Gilbert F 1960 *J. Acoustic Soc. Am.* **32** 841
[7] Gilbert F and Knopoff L 1959 *J. Acoust. Soc. Am.* **31** 1169
[8] Jeffreys H and Jeffreys B S 1956 *Methods of Mathematical Physics* (Cambridge)
[9] Jha R 1974 *Proc. Cambridge Philos. Soc.* **75** 391
[10] Keller J B 1958 *Div. Electromag. Res. Inst. Math. Sci. New York Univ. Res. Rep.* EM -115
[11] Mishra S K 1964a *Proc. Cambridge Philos. Soc.* **60** 295
[12] Mishra S K 1964b *Proc. Indian Acad. Sci.* **A59** 21
[13] Rajhans B K and Mishra S K 1980 *Proc. Indian Acad. Sci. (Math. Sci.)* **89** 3
[14] White J E 1965 *Seismic waves. Radiation, transmission and attenuation* (New York : McGraw-Hill)

Proc. Indian Acad. Sci. (Math. Sci.), Vol. 91, Number 2, July 1982, pp. 155-165.

The wall jet flow of a conducting gas over a permeable surfa[ce in] the presence of a variable transverse magnetic field

J L BANSAL, M L GUPTA and S S TAK
Department of Mathematics, University of Jodhpur, Jodhpur, India

MS received 8 May 1981; revised 10 February 1982

Abstract. The effects of the magnetic field, Mach number and the per[meability] parameter on the wall jet flow (radial or plane) of an electrically conduc[ting gas] spreading over a permeable surface have been investigated. Taking the [Prandtl] number of the fluid as unity and assuming a linear relationship between [viscosity] and temperature, it is found that similar solutions for the velocity dis[tribution] exist for a specified distribution of the normal velocity along the wall [and the] corresponding distribution of the transverse magnetic field. Previous non-[magnetic] flow results have been improved by adopting a new and simple transform[ation of] variables.

Keywords. Boundary layer theory; magnetohydrodynamics; jet flow.

1. Introduction

The non-magnetic wall jet in a gas medium has been studied by means of [similar] solutions [7, 4, 9, 5, 2, 3]. Fox and Steiger [5] studied the fluid d[ynamic] characteristics of a wall jet of an electrically non-conducting gas spread[ing] over a permeable plane surface in otherwise stationary ambient surround[ings by] posing an eigenvalue problem on the velocity field. The problem was [solved] numerically for various values of suction/injection parameter. They also o[btained] a particular solution of the temperature distribution assuming the Prandtl [number] of the fluid as unity. Bansal [1] reconsidered the work of Fox and Ste[iger] and extended it to the study of thermal boundary layer for arbitrary va[lues of] the Prandtl number. He found that similar solutions exist both for rad[ial and] plane wall jets, for a specified distribution of normal velocity along the [wall].

In the present paper we have studied the effects of the magnetic field[, Mach] number and the permeability parameter on the wall jet (radial or plane) [of an] electrically conducting gas spreading over a permeable plane wall. Keep[ing the] works of Fox and Steiger [5] and Bansal [1] in view, we have obtai[ned] similar solution for the velocity field for various values of the suction/i[njection] parameter α and magnetic parameter m. A linear relationship between v[iscosity] and temperature has been assumed and the Prandtl number of the fluid i[s taken] as unity. It is found that the similar solution exists for a specified distr[ibution]

of the normal velocity along the wall and the corresponding distribution of the transverse magnetic field. The improved values of the results of Fox and Steiger have been obtained as a particular case.

One of the important conclusions of the present study is that the coefficient of skin-friction is independent of the Mach number but decreases with the increase in the value of the magnetic parameter m. It is noted that the effect of injection is to reduce the skin friction coefficient at all stations downstream as compared to the value for impermeable wall, whereas in the case of suction it first decreases and then increases monotonically downstream both for radial and plane jets.

2. Formulation of the problem

Let an electrically conducting gas at a temperature T_∞ be discharged through a small orifice (slit or circular) spreading out over a permeable plane surface in the presence of a variable transverse magnetic field and mix with an ambient gas being initially at rest having a temperature T_∞, so that the thermal boundary layer is formed only due to viscous dissipation and Joule heating.

Taking the origin in the orifice and the coordinate axes x and y along and normal to the plane wall respectively, the equation of state, continuity and the boundary layer approximations to the momentum and the energy equations are:

State :

$$p = \rho RT, \tag{1}$$

Continuity :

$$\frac{\partial}{\partial x}(\rho x^i u) + \frac{\partial}{\partial y}(\rho x^i v) = 0, \tag{2}$$

Momentum :

$$\rho u \frac{\partial u}{\partial x} + \rho v \frac{\partial u}{\partial y} = \frac{\partial}{\partial y}\left(\mu \frac{\partial u}{\partial y}\right) - \sigma_e B^2 u, \tag{3}$$

Energy :

$$\rho u \frac{\partial \bar{T}}{\partial x} + \rho v \frac{\partial \bar{T}}{\partial y} = \frac{1}{\sigma}\frac{\partial}{\partial y}\left(\mu \frac{\partial \bar{T}}{\partial y}\right) + \frac{\mu}{c_p}\left(\frac{\partial u}{\partial y}\right)^2 + \frac{\sigma_e B^2}{c_p} u^2, \tag{4}$$

where

$$\bar{T} = T - T_\infty, \quad \sigma = \frac{\mu c_p}{k} \text{ (Prandtl number)}, \tag{5}$$

The boundary conditions are as follows:

$$y = 0: \quad u = 0, \quad v = v_w(x); \quad \bar{T} = 0 \quad \text{(isothermal wall)},$$

$$\text{or} \quad \frac{\partial \bar{T}}{\partial y} = 0 \text{ (adiabatic wall)}, \tag{6}$$

$$y = \infty : \frac{\partial u}{\partial y} = 0, \frac{\partial u}{\partial x} \to -\frac{\sigma_e B^2}{\rho},$$

$$\frac{\partial \bar{T}}{\partial y} = 0, \quad \frac{\partial \bar{T}}{\partial x} \to \frac{\sigma_e B^2}{\rho c_p} u; \tag{7}$$

where for injection $v_w(x) > 0$ and for suction $v_w(x) < 0$.

By the boundary layer approximation the pressure is uniform everywhere and hence the equation of state (1) implies that

$$\rho T = \text{constant} = \rho_\infty T_\infty. \tag{8}$$

3. Analysis

Introducing the stream function ψ, such that

$$\rho x^t u = \frac{\partial \psi}{\partial y} \text{ and } \rho x^t v = -\frac{\partial \psi}{\partial x}. \tag{9}$$

The solution of the coupled partial differential equations of § 2 is facilitated by taking a linear relationship between viscosity μ and temperature T, i.e.,

$$\frac{\mu}{\mu_\infty} = \frac{T}{T_\infty}. \tag{10}$$

Hence,

$$\rho \mu = \text{constant} = \rho_\infty \mu_\infty. \tag{11}$$

Following Levy–Lees [8], let us transform the independent variables (x, y) to (ξ, η), such that

$$\xi = \int_0^x \mu_\infty \rho_\infty u_x x^{2t} dx, \tag{12}$$

$$\eta = \frac{\rho_\infty u_x x^t}{\sqrt{2\xi}} \int_0^y \frac{\rho}{\rho_\infty} dy, \tag{13}$$

where the subscript x represents some reference value, yet to be prescribed, of the velocity component u in the boundary layer.

Taking the following forms of stream function and the temperature difference $\bar{T}$:

$$\psi = (2\xi)^{1/2} f_m(\eta), \tag{14}$$

$$\bar{T} = -\frac{u_x^2}{2c_p} h_m(\eta), \tag{15}$$

the momentum and energy equations reduce to

$$f_m''' + f_m f_m'' - 2\alpha f_m'^2 - 2\beta f_m' = 0, \tag{16}$$

and

$$h_m'' + \sigma(f_m h_m' - 4\alpha f_m' h_m) = 2\sigma(f_m''^2 + 2\beta f_m'^2), \tag{17}$$

respectively, with the boundary conditions

$$\eta = 0:\ f_m = f_m(0)\ \text{(say)},\ f'_m = 0;\ \ h_m = 0\ \ \text{(isothermal wall)},$$
$$\text{or } h'_m = 0\ \text{(adiabatic wall)}, \tag{18}$$
$$\eta \to \infty:\ f'_m \to -\frac{\beta}{\alpha};\ h_m \to \frac{\beta^2}{\alpha^2}$$

where a prime denotes differentiation with respect to η and α, β are given by

$$\alpha = \frac{\xi}{u_e}\frac{du_e}{d\xi}\ \text{(wall permeability parameter)}, \tag{19}$$

and

$$\beta = \left(\frac{\xi}{\rho_\infty \mu_\infty x^{2i} u_s^2}\right)\frac{\sigma_e B^2}{\rho}. \tag{20}$$

The requirement of similar solution is that α and β must be independent of x. Taking α as constant, from (19) and (12) it follows that

$$\frac{u_s}{u_{sc}} = \left(\frac{x}{x_e}\right)^{\alpha(2i+1)/1-\alpha},\ \alpha \neq 1. \tag{21}$$

Hence,

$$v_w(x) = -A(1-\alpha)^{-1/2} f_m(0)\left(\frac{x}{x_c}\right)^n, \tag{22}$$

where

$$A = \mu_w\left\{\frac{(2i+1)u_{sc}}{2\rho_\infty \mu_\infty x_c}\right\}^{1/2}\ \text{and}\ n = \frac{2\alpha(i+1)-1}{2(1-\alpha)}. \tag{23}$$

Therefore, $f_m(0) < 0$ (injection) and $f_m(0) > 0$ (suction). For β to be independent of x, B is assumed to be of the following form:

$$B = B_0\, x^{(2\alpha+2i\alpha-1)/2(1-\alpha)},\ \ \alpha \neq 1. \tag{24}$$

Hence,

$$\beta = \frac{m(1-\alpha)\rho_\infty}{\rho}, \tag{25}$$

where

$$m = \frac{x_c^{\alpha(2i+1)/(1-\alpha)}}{(2i+1)u_{sc}}\frac{\sigma_e B_0^2}{\rho_\infty}\ \text{(magnetic interaction parameter)}. \tag{26}$$

A useful simplification is made by taking the Prandtl number of the conducting gas as unity. In such a case, it may easily be seen that

$$h_m = f'^2_m,\ \sigma = 1, \tag{27}$$

is a solution of the energy equation (17) both for isothermal and adiabatic walls. It is, therefore, required to solve the momentum equation only.

Hence,

$$\frac{T}{T_\infty} = 1 - \frac{u_s^2}{2c_p T_\infty} f'^2_m, \tag{28}$$

and

$$\frac{\rho_\infty}{\rho} = 1 - \frac{1}{2}(\gamma - 1)\, Ma^2 f_m'^2, \tag{29}$$

where the relationship

$$u_s^2 = (\gamma - 1)\, Ma^2\, c_p T_\infty, \tag{30}$$

has been used. In general, Mach number is variable and shall violate the similarity requirements. However, in the present solution it has been taken as constant, an assumption which is usually made in the study of compressible boundary layers [10].

Therefore, the momentum equation (16), in view of (25) and (29), becomes

$$f_m''' + f_m f_m'' - 2\alpha f_m'^2 - 2m(1-\alpha)\left\{1 - \frac{(\gamma-1)}{2} Ma^2 f_m'^2\right\} f_m' = 0, \tag{31}$$

with the boundary conditions

$$\eta = 0 \quad : f_m = f_m(0),\ f_m' = 0,$$

$$\eta = \infty : \left[\frac{m(1-\alpha)}{\alpha}\left\{1 - \frac{(\gamma-1)}{2} Ma^2 f_m'^2\right\} + f_m'\right] \to 0. \tag{32}$$

In adiditon, the two-point boundary value problem posed by (32) implies the existence of an eigenvalue problem that is, for each value of α only one solution $f_m(\eta)$ with the correct asymptotic behaviour as $\eta \to \infty$ can be found. Therefore, each case of suction or injection is related to the eigenvalue α and requires a specific form of the reference velocity u_s and the normal velocity $v_w(x)$ along the wall for a given magnetic field. This means that α and $f_m(0)$ cannot be prescribed simultaneously. In fact one determines the other and it will be easy to prescribe α and determine $f_m(0)$, with the help of Lagrange's three-point interpolation formula, such that the condition at infinity is achieved.

Before we presctibe α, let us determine its range which is obtained by the following two expressions:

$$f_0(0) f_0''(0) = -2(\alpha + 1)\int_0^\infty f_0 f_0'^2\, d\eta, \tag{33}$$

$$f_0''(0) = -(2\alpha + 1)\int_0^\infty f_0'^2\, d\eta. \tag{34}$$

Equation (33) is obtained by multiplying (31) with f_0 and then integrating with respect to η between the limits 0 to ∞ and putting $m = 0$; whereas (34) is obtained by direct integration of (31) between the same limits and taking $m = 0$.

(i) For impermeable wall, since $f_0(0) = 0$, relation (33) gives $\alpha = -1$ and $\int_0^\infty f_0 f_0'^2\, d\eta = \text{constant}$, which is the same integral condition as was obtained by Akatnow and Glauert [6].

(ii) For $f_0''(0) = 0$ (blow-off condition), it is concluded from (34) that

$$\alpha = -\tfrac{1}{2} \text{ and } \int_0^\infty f_0'^2\, d\eta = \text{constant},$$

Further, (33) gives

$$\int_0^\infty f_0 f_0'^2 \, d\eta = 0.$$

Since $f_0'^2$ is always positive, it is implied in this case that f_0 must be asymmetric function varying from $f_0 = f_0(0) < 0$ to $f_0 = f_0(\infty) = 1$ as η varies from 0 to ∞. Hence, $\alpha = \frac{1}{2}$ implies injection at the permeable wall.

(iii) For $f_0''(0) < 0$ (region of back flow), since $\int_0^\infty f_0'^2 \, d\eta$ is always positive, it is concluded from (34) that $\alpha > -\frac{1}{2}$. As the boundary layer approximations breakdown in the region of back-flow, we shall limit ourselves to the value of $\alpha < -\frac{1}{2}$.

(iv) For $f_0''(0) > 0$, from (i), (ii) and (iii) it may be concluded that

$$-1 < \alpha \leqslant -\tfrac{1}{2} \text{ for } f_0(0) < 0 \text{ (injection)},$$

$$\alpha = -1 \text{ for } f_0(0) = 0 \text{ (impermeable wall)},$$

$$-\infty < \alpha < -1 \text{ for } f_0(0) > 0 \text{ (suction)}.$$

Finally, the range of α is

$$-\infty < \alpha \leqslant -\tfrac{1}{2}. \tag{35}$$

4. Transformation of variables

To simplify the analysis let us make the following transformation of variables :

$$f_m(\eta) = \{a\alpha(am - 1)\}^{-1/2} F_m(\zeta) \tag{36}$$

and

$$\zeta = a^{-1/2} \{\alpha(am - 1)\}^{1/2} \eta. \tag{37}$$

Hence,

$$f_m'(\eta) = a^{-1} F_m'(\zeta), \tag{38}$$

$$f_m''(\eta) = a^{-3/2} \{\alpha(am - 1)\}^{1/2} F_m''(\zeta), \tag{39}$$

and (31) reduces to

$$\alpha(am - 1) F_m''' + F_m F_m'' - 2\alpha F_m'^2$$

$$- 2am(1 - \alpha)\left\{1 - \frac{(\gamma - 1) Ma^2}{2a^2} F_m'^2\right\} F_m' = 0, \tag{40}$$

with the boundary conditions

$$\zeta = 0 : F_m = \{a\alpha(am - 1)\}^{1/2} f_m(0), \; F_m' = 0, \; F_m'' = 1 \quad m > 0$$

$$\zeta = \infty : F_0(\infty) = (-a\alpha)^{1/2} ;$$

$$\left[\frac{am(1 - \alpha)}{\alpha}\left\{1 - \frac{(\gamma - 1) Ma^2}{2a^2} F_m'^2\right\} + F_m'\right] \to 0, \tag{41}$$

where a prime denotes differentiation with respect to ζ,

Here we have taken $F_m''(0) = 1$, which is permissible without loss of generality, due to the presence of a free constant a in (39).

4.1. *Determination of a*

The value of a, which is independent of m, can be easily obtained by taking $m = 0$ in (40) and (41).

The corresponding equation was solved by Fox and Steiger [5] by a very tedious method. A simple and accurate method, which we suggest, is the following: Prescribe a value of α and interpolate the value of $F_0(0)$ such that $F_0'(\infty) = 0$. This will give us a corresponding value of $F_0(\infty)$, which in turn will determine a, $f_0(0)$ and $f_0''(0)$ as follows:

$$a^{1/2} = (-\alpha)^{-1/2} F_0(\infty), \tag{42}$$

$$f_0(0) = F_0(0)/F_0(\infty), \tag{43}$$

$$f_0''(0) = \alpha^2/F_0^3(\infty). \tag{44}$$

The calculated values of a, $f_0(0)$ and $f_0''(0)$ for various values of α are given in table 1. One may see that this method gives more accurate results.

Now, we may think of the integration of equation (40) in the following manner: Prescribe the values of γ, Ma and m $(am < 1)$. For a chosen value of α, take the corresponding value of a from table 1 and interpolate $F_m(0)$ so that the given condition at infinity is achieved. The value of $f_m(0)$ is, now, given by

$$f_m(0) = \{a\alpha(am - 1)\}^{-1/2} F_m(0), \tag{45}$$

and

$$f_m''(0) = a^{-3/2}\{a(am - 1)\}^{1/2}. \tag{46}$$

It may be noted here that $f_m''(0)$ is independent of Mach number; however, $f_m(0)$ will very much depend on it.

Table 1. Solution of the equation (40), $m = 0$ (initial values).

	α	a	$f_0(0)$ Present method	$f_0(0)$ Fox and Steiger [5]	$f_0''(0)$ Present method	$f_0''(0)$ Fox and Steiger [5]
Injection	−0·55	5·50806	−0·76139	−0·749	0·0948	0·061
	−0·60	3·87370	−0·61563	−0·591	0·10160	0·106
	−0·70	3·02650	−0·38746	−0·379	0·15890	0·159
	−0·75	2·88975	−0·29793	−0·298	0·1763	0·186
Impermeable wall	−1·0	2·72575	0·0	0·0	0·2222	0·222
Suction	−1·25	2·83350	0·17982	0·193	0·2344	0·235
	−1·5	3·03178	0·30500	0·305	0·23200	0·232
	−2·0	3·55905	0·46538	0·472	0·2106	0·208
	−5·0	7·24786	0·77604	0·776	0·1146	0·117
	−10·0	13·55271	0·88592	0·892	0·06338	0·0688
	−50·0	57·36662	0·9771	0·976	0·01627	0·0235

5. Transformation to the physical plane

Inverting the variables, we find that

$$y = \frac{\sqrt{2a\xi}}{\rho_\infty u_e x^i \sqrt{a(am-1)}}\left[\zeta - \frac{(\gamma-1)Ma^2}{2a^2}\int_0^\zeta F_m'^2 \, d\zeta\right] \tag{47}$$

and

$$\xi = \rho_\infty \mu_\infty u_e (1-a)\frac{x^{2i+1}}{(2i+1)}. \tag{48}$$

Thus we see that the effect of the index i (i.e., the nature of the orifice) is not felt in the transformed plane, it comes into play only when we switch over to the physical plane.

6. Coefficient of skin-friction

The shearing stress τ_w is given by

$$\tau_w = \left(\mu \frac{\partial u}{\partial y}\right)_{y=0} = \frac{\mu_w \rho_w u_e^2 x^i}{\sqrt{2\xi}} f_m''(0). \tag{49}$$

Hence, the coefficient of skin-friction is obtained as

$$C_{f_m}(m,a) = \frac{\tau_w}{\frac{1}{2}\rho_\infty u_e^2}$$

$$= \left\{\frac{2(2i+1)\mu_\infty}{x_0 u_{e_0} \rho_\infty}\right\}^{1/2}\left(\frac{x}{x_0}\right)^{-(2ia+1)/2(1-a)}(1-a)^{-1/2} f_m''(0). \tag{50}$$

For impermeable wall ($a = -1$) and in the absence of magnetic field ($m = 0$), we have [6] :

$$[f_0''(0)]_{a=-1} = \frac{2}{9}. \tag{51}$$

Hence,

$$\frac{C_{f_m}(m,a)}{c_{f_0}(0,-1)} = \frac{9}{\sqrt{2}}\left(\frac{x}{x_0}\right)^{-(2i+1)(1+a)/4(1-a)}(1-a)^{-1/2} f_m''(0). \tag{52}$$

7. Numerical discussion

The velocity distribution, in the absence of magnetic field (i.e., $m = 0$) and for various values of the permeability parameter a, is plotted against the transformed similarity variable ζ in figure 1. It is noted that the effect of fluid injection is to increase the velocity in a considerable neighbourhood of the wall and to decrease it thereafter, resulting in the decrease of the width of the jet. A reverse phenomenon happens in the case of suction.

Taking $\gamma = 1{\cdot}4$ and $m = 0{\cdot}15$, the velocity distribution in a hydromagnetic plane wall jet spreading over a permeable surface is plotted against the transformed similarity variable ζ in figure 2, for two values of Mach number 0 and 2

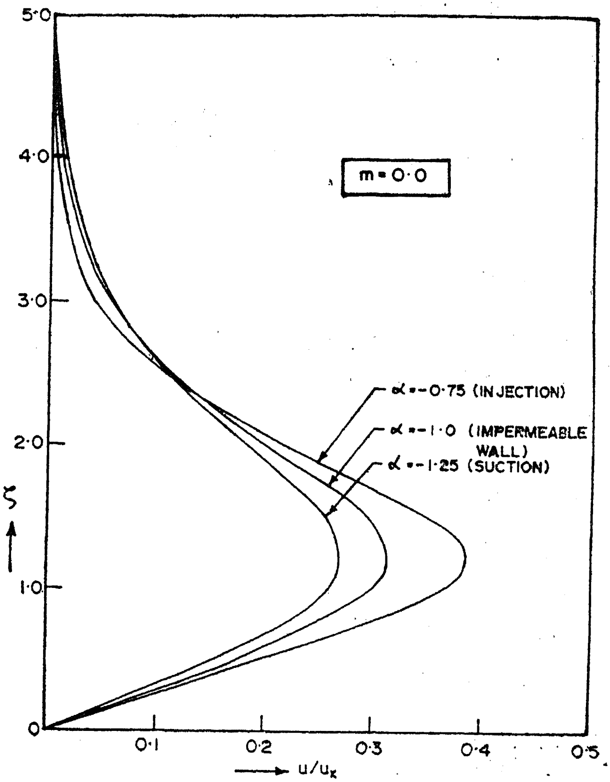


Figure 1. Velocity distribution in a wall jet of a gas spreading over a permeable surface for various values of α.

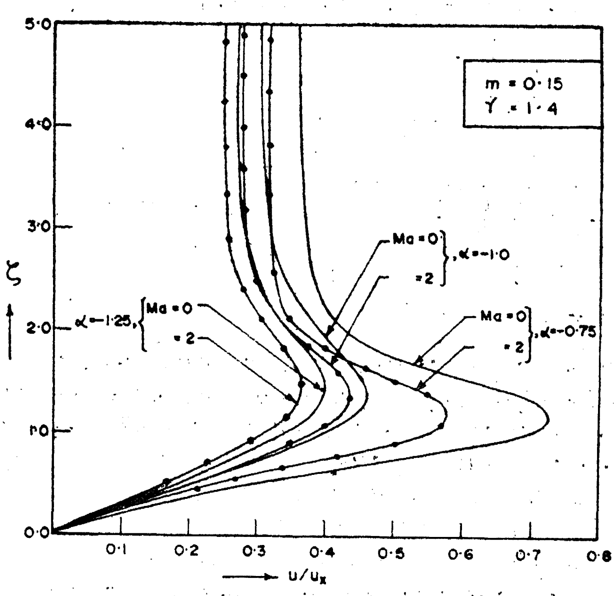


Figure 2. Velocity distribution in a hydromagnetic wall jet of a conducting gas spreading over a permeable surface for various values of $\alpha \cdot m = 0{\cdot}15$ and $Ma = 0, 2$.

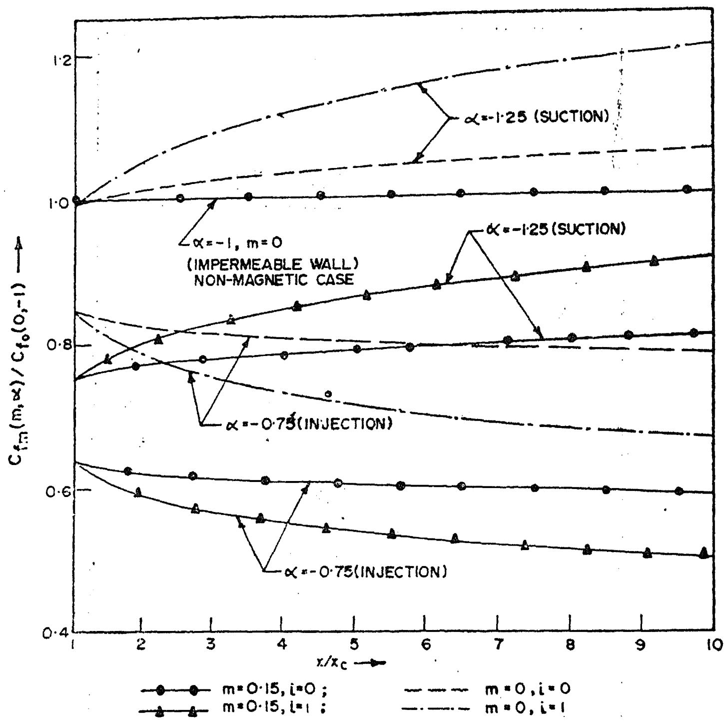


Figure 3. Variation of relative coefficient of skin-friction in a hydromagnetic wall jet of a conducting gas spreading over a permeable surface for various values of α and m and for $i = 0, 1$.

and for three values of the wall permeability parameter $\alpha = -0\cdot75$, -1 and $-1\cdot25$. It is found that the effect of the Mach number is to decrease the velocity at all points and in all cases of the permeability of the wall. As far as the question of the permeability of the wall is concerned, its effect is of the same nature as we observed in figure 1. The effect of the magnetic field can be seen by combining figures 1 and 2 and it is concluded that the magnetic field contributes towards the increase of the width of the jet.

One of the important studies of the present analysis is the variation of the relative coefficient of skin-friction, which is plotted in figure 3 against the dimensionless distance x/x_c down the stream. It is found to be independent of Mach number, an outcome of the linear relationship between viscosity and temperature, but decreases with the increase in the value of the magnetic parameter m. It is noted that the effect of gas injection is to reduce the skin-friction coefficient at all stations downstream as compared to the value for impermeable wall, whereas in the case of suction it first decreases and then increases monotonically downstream both for radial and plane jets.

Nomenclature

$B(x)$ strength of variable transverse magnetic field
c_p specific heat at constant pressure

$C_{fm}(m, \alpha)$	coefficient of skin-friction
f, F	velocity distribution functions introduced in (14) and (36) respectively
h	temperature distribution function
i	0 or 1 for plane or radial wall jet respectively
m	magnetic interaction parameter defined by (26)
Ma	Mach number
p	pressure
T	temperature in the boundary layer
u, v	velocity components along and normal to the plane wall respectively
x, y	coordinates along and normal to the plane wall

Greek symbols

α	suction/injection parameter
β	defined by (20)
γ	ratio of the two specific heats c_p and c_v
ζ	similarity variable defined by (37)
ξ, η	Levy–Lees variables defined in (12) and (13)
k	thermal conductivity
μ	coefficient of viscosity
ρ	density
σ	Prandtl number
σ_e	electrical conductivity
τ_ω	shearing stress at the wall
ψ	stream function

Subscripts

c	refers to values at a reference station in the boundary layer
o, m	refers to non-magnetic and magnetic cases respectively
w	refers to values at the wall
x	denotes reference value in the boundary layer
∞	refers to values at the edge of the boundary layer.

References

[1] Bansal J L 1981 A new approach to some hydrodynamic and hydromagnetic boundary layer flow problems. D.Sc. Thesis, Univ. of Jodhpur, Jodhpur

[2] Bansal J L and Tak S S 1978 *ZAMP* **29** 742–756

[3] Bansal J L and Tak S S 1979 *Indian J. Pure Appl. Math.* **10** (12) 1469–1483

[4] Bloom M H and Steiger M H 1958 *Proc. Third U.S. Nat. Congr. Appl. Mech.* p. 717 Brown University

[5] Fox H and Steiger M H 1963 *J. Fluid Mech.* **15** 4

[6] Glauert M B 1956 *J. Fluid Mech.* **1** (6) 625–643

[7] Glauert M B 1957 Boundary Layer Research Symp. Freiburg/Br.

[8] Lees L 1956 *Jet. Prop.* **26** 259–269

[9] Riley N 1958 *J. Fluid Mech.* **4** (6) 615–628

[10] White Frank M 1974 *Viscous fluid flow* (USA; McGraw-Hill) pp. 589–592

Proc. Indian Acad. Sci. (Math. Sci.), Vol. 91, Number 3, November 1982, pp. 167–182.

On invariant convex cones in simple Lie algebras

S KUMARESAN and AKHIL RANJAN
School of Mathematics, Tata Institute of Fundamental Research, Homi Bhabha Road, Bombay 400 005, India

MS received 19 December 1981 ; revised 22 January 1982

Abstract. This paper is devoted to a study and classification of G-invariant convex cones in g, where G is a lie group and g its Lie algebra which is simple. It is proved that any such cone is characterized by its intersection with h–a fixed compact Cartan subalgebra which exists by the very virtue of existence of proper G-invariant cones. In fact the pair (g,k) is necessarily Hermitian symmetric.

Keywords. Lie algebra ; adjoint group ; invariant convex cones.

1. Introduction

Let g be a real simple Lie algebra. Let G be its adjoint group. A closed convex cone in g invariant under the adjoint action of G will be referred to simply as invariant cone. The study of such cones was begun by Vinberg [9]. He proved (loc. cit) that for g to have an invariant cone it is necessary and sufficient that G is the group of isometries of a Hermitian symmetric space. Hence we shall always assume that this is the case. In his paper [9] Vinberg poses some problems. We answer two of his questions in this paper.

One question (problem 4 in [9]) of Vinberg asks for characterisation of invariant convex cones *via* their intersection with the compact Cartan subalgebra. Vinberg also asks whether it is true that for an invariant convex cone Q and a compact Cartan subalgebra h of g, $(Q \cap h)^* = (Q^* \cap h)$. Here Q^* denotes the dual of Q. (For definition and notation, see the main part of the paper.) As an answer to this question we prove that an invariant convex cone is completely determined by its intersection with a compact Cartan subalgebra and the answer to the second part of the question is yes.

To state this more precisely, we need the standard notation such as $g = k \oplus p$, W_K the Weyl group of (k, h), etc., and two W_K-invariant cones $C_{\min}(g) \cap h$ and $C_{\max}(g) \cap h$. (The last two objects are defined by Vinberg and for their definition, refer to § 2). Let p_h stand for the projection of g onto h *via* the Killing form. With this notation, we can state our main result.

Main result : If Q is an invariant convex cone in g, then Q is the invariant convex cone generated by $p_h(Q)*Q \cap h$. Conversely if C is a W_K-invariant cone

such that $C_{\min}(g) \cap h \subseteq C \subseteq C_{\max}(g) \cap h$, then the invariant convex cone $Q(C)$ generated by C in g is such that $Q(C) \cap h = C = p_h(Q(C))$.

From this it trivially follows that whenever $C_{\min}(g) \neq C_{\max}(g)$, there exists a continuum of invariant cones. This answers, in particular, a question of Vinberg (problem 1 [9]).

Again from the main result, we can deduce for any invariant convex cone Q, we have $(Q \cap h)^* = Q^* \cap h$. In particular, in view of the main result, Q is self-dual if and only if $Q \cap h$ is self-dual in h. Hence to show that self-dual cones exist in g, one has only to construct self-dual cones in (the Euclidean space) h. Using our description of $C_{\max}(g) \cap h$ (Prop. 4) and $C_{\min}(g) \cap h$, one can easily see that self-dual W_K-invariant cones can be squeezed between $C_{\min}(g) \cap h$ and $C_{\max}(g) \cap h$. Hence the existence of self-dual invariant convex cones in g follows.

Now the natural question arises: Are there only finitely many self-dual invariant convex cones in g?

Again the answer is No! In SO* (6) ($\sim$ SU (3, 1)) we show that there exists an infinity of self-dual cones. Of course, by what we said above the construction boils down to that of S_3-invariant cones in R^3 squeezed in between two cones (which are to be specified). The details of this last construction are given in the appendix.

(We thank M S Raghunathan for pointing out a gap in the proof of Theorem 11 in the first version.)

Proof of the main result

Let G be a connected real simple Lie group and g its Lie algebra. Let θ be a Cartan involution of g and let $g = k \oplus p$ be the corresponding Cartan decomposition. Let K be the connected subgroup whose Lie algebra is k. We wish to study closed convex cones in g which are invariant under the adjoint action of G. g admits invariant cones if and only if there exists a K-invariant vector in g under the adjoint action [9]. This implies that g will have invariant cones if and only if k has a (one-dimensional) centre. Let k_0 be a (non-zero) element of k so that the centre of k is $Z(k) = R k_0$. It is a well-known fact that k_0 induces a complex structure on p *via* the adjoint action [6: p. 375]. Let J denote this complex structure on p.

Now let g_C, k_C, p_C denote the complexifications of k and p respectively. Let h be a maximal abelian subalgebra of k. Then h_C is a Cartan subalgebra of g_C. Note that $k_0 \in h$. Let $\triangle$ denote the set of roots of (g_C, h_C). We choose a positive system P for $\triangle$ compatible with the complex structure on G/K. By this we mean a positive system in which every positive non-compact root (*i.e.*, a root $\alpha \in P$ such that $g^\alpha \subseteq p_C$) is greater than every compact root (*i.e.*, a root α with $g^\alpha \subseteq k_C$). We let P_k (resp. P_n) denote the set of positive compact (resp. non-compact) roots in P. Then $P = P_k \cup P_n$. Then $p^+ = \sum_{+\alpha \in P_n} g^\alpha$ is an abelian subalgebra invariant under the adjoint action of k_C. Let (p, J) denote the complex vector space whose underlying real vector space is p with the complex structure J. It is well-known that (p, J) is an irreducible k_C module and (p, J) and p^+ are equivalent as k_C-modules. For what follows it is important to understand this equivalence in an explicit manner which we describe now.

Let $0 \neq Z_{\pm\alpha} \, \alpha \in g^{\pm\alpha}$ for $\in P_n$. Set $X_\alpha = Z_\alpha + Z_{-\alpha}$ and $Y_\alpha = i(Z_\alpha - Z_{-\alpha})$. Then $X_\alpha, Y_\alpha \in p$ and are such that $JX_\alpha = Y_\alpha$ and $JY_\alpha = -X_\alpha$. These follow from trivial computations and the fact that all roots are purely imaginary on the real space h. It can also be easily checked that for $H_\alpha = [Z_\alpha, Z_{-\alpha}]$, we have

(i) $[iH_\alpha, X_\alpha] = 2Y_\alpha$,

(ii) $[iH_\alpha, Y_\alpha] = -2X_\alpha$ and

(iii) $[X_\alpha, Y_\alpha] = -2i H_\alpha$.

Then the map $X \mapsto X - iJX$ is a k_C-intertwining map of the modules (p, J) and p^+. Under this map $X_\alpha \mapsto 2Z_\alpha$ and $Y_\alpha \mapsto 2iZ_\alpha$.

Let $C_{\min}(g) = \text{Con } G \cdot k_0$, the smallest closed convex cone containing k_0. Here $G \cdot k_0$ stands for $\{Ad g \cdot k_0 : g \in G\}$. For the coadjoint representation of G on g' we can similarly define $C_{\min}(g')$. Then $C_{\max}(g)$ is defined to be the cone $\{X \in g : (X, Y) \geqslant 0 \text{ for } Y \in C_{\min}(g')\}$. Let $-B(,)$ denote the negative of the Killing form of g. Identifying g' with g *via* this non-degenerate form, we can write

$$C_{\max}(g) = \{X \in g: \; -B(X, Y) \geqslant 0 \;\text{ for }\; Y \in C_{\min}(g)\}.$$

Note that $k_0 \in C_{\max}(g)$ and this is the reason why we used $-B(,)$ to identify g, with g. Then $C_{\max}(g)$ is the unique maximal invariant cone in g : unique, to within multiplication by ± 1 [9 ; theorem 3].

Our main aim is to give a description of $Q \cap h$ for an invariant cone Q in g. Let W_K denote the Weyl group of the pair (k, h).

Observation : If Q is an Ad (G)-invariant cone in g, then $Q \cap h$ is a W_K-invariant cone in h.

That $Q \cap h$ is a cone in the real vector space h is easily seen. That it is W_K-invariant follows from the fact that given any $w \in W_K$ there exists $k \in K$ such that Ad k, restricted to h, induces w. This suggests that we should investigate what happens when we start with a W_K-invariant cone C in h and consider the smallest Ad (G)-invariant closed convex cone in g containing C. As a first step in this investigation we prove

Lemma 2 : Let C be a W_K-invariant cone in h. Then Ad $K(C)$ is a cone in k. Note that we are not saying that the cone generated by

$$\{\text{Ad}\, k(C) : k \in K\} \text{ in } k \,!$$

To prove this lemma we make use of a convexity theorem (or rather its infinitesimal analogue) due to Kostant [6]. To state this theorem we need some notation, which we now set up. Let G be any connected semisimple Lie group with finite centre. Let g be its Lie algebra. Let $g = k \oplus p$ be a Cartan decomposition of g. Fix a maximal abelian subspace a of p. Let p_a denote the orthogonal projection of p onto a with respect to the Killing form. Let W_a denote the little Weyl group of (g, a). Recall that W_a is defined to be the normalizer of a in K (the connected subgroup corresponding to k) modulo the centralizer of a in K.

Theorem [6]: Let $H \in a$. Then $p_a(\text{Ad}\, K(H))$ is the convex-hull of $\{w \cdot H : w \in W_a\}$.

A Morse-theoretic proof of this (infinitesimal) version can be found in [4; theorem 1, Ch. 1; see also 2].

Let now U be any connected compact group with Lie algebra u. Let h be a Cartan subalgebra in u. Let W denote the Weyl group of (u, h). Let $u_s = [u, u]$ denote the semisimple part of u. Let z be the centre of u. Then $h_s = h \cap u_s$ is a Cartan subalgebra of u_s and we have $h = z \oplus h_s$, and $W(u_s, h_s) = W(u, h)$. We now apply the above result of Kostant to $g = u_s + iu_s$, and $a = ih_s$ $(i = \sqrt{-1})$. It is easily verified that W_a is canonically isomorphic to W. Hence we deduce that if p_{h_s} denotes the projection of u_s onto h_s with respect to the Killing form then $p_{h_s}(\mathrm{Ad}\, U_s(H))$ is the convex-hull of $\{wH : w \in W\}$ for any $H \in h_s$. If $H \in h$ and $u \in U$, then we write $H = Z + H_s, Z \in z$, $H_s \in h_s$ and $\mathrm{Ad}\, u = \mathrm{Ad}\, z\, \mathrm{Ad}\, u_s$ with z in the centre of U and u_s in the semisimple part of U. Then we see that $\mathrm{Ad}\, u\, (H) = Z + \mathrm{Ad}\, u_s\, (H_s)$ and hence $p_h(\mathrm{Ad}\, U(H))$ is the convex-hull of $\{wH ; w \in W\}$ in the general case too.

Proof of Lemma 2 : Since K is compact, and C is closed in h, $\mathrm{Ad}\, K(C)$ is closed in k. Also, since C is a cone, for any $t > 0$, $k \in K$ and $H \in C$, $t \cdot \mathrm{Ad}\, k\, (H) = \mathrm{Ad}\, k\, (tH)$. Hence it is enough to show that for any $X, Y \in \mathrm{Ad}\, K(C)$ we have $X + Y \in \mathrm{Ad}\, K(C)$. Choose $k \in K$ so that $\mathrm{Ad}\, k\, (X + Y) \in h$. This is possible since h is maximal abelian in k. If $X = \mathrm{Ad}\, k_1 H_1$, $Y = \mathrm{Ad}\, k_2 H_2$ for some $k_1, k_2 \in K$ and $H_1, H_2 \in C$, then $\mathrm{Ad}\, k\, (X + Y) = H$ for some $H \in h$ implies that $\mathrm{Ad}\, k\, k_1 H_1 + \mathrm{Ad}\, k\, k_2 H_2 = H$. Since by what we have seen above $p_h(\mathrm{Ad}\, kk_i\, H_i)) \in$ convex-hull of $\{w \cdot H_i ; w \in W_K\}$ and since C is a W_k-invariant convex cone in h, it follows that $X + Y \in \mathrm{Ad}\, K(C)$. This proves the lemma.

Let p_h also denote the projection of g onto h *via* the Killing form.

Corollary 3 : The following are equivalent for any two $\mathrm{Ad}\,(K)$-invariant cones Q_1 and Q_2 in k:

(i) Two $\mathrm{Ad}\,(K)$-invariant (closed, convex) cones in k are equal: $Q_1 = Q_2$.

(ii) $Q_1 \cap h = Q_2 \cap h$.

(iii) $Q_1 \cap h = P_h(Q_1) = p_h(Q_2) = Q_2 \cap h$.

Proof: (i) $\Rightarrow$ (ii) : This is trivial.

(ii) $\Rightarrow$ (i). Since $Q_1 \cap h$ is a W_K-invariant convex cone in h, $Q_i \supseteq \mathrm{Ad}\, K(Q_i \cap h)$, the latter being a cone in k by lemma 2. If $X \in Q_i$, then $\mathrm{Ad}\, k\, X \in h$ for some $k \in K$ and since Q_i is $\mathrm{Ad}\, K$-invariant, $\mathrm{Ad}\, k\, X \in Q_i \cap h$. Hence $X \in \mathrm{Ad}\, k^{-1}(Q_i \cap h)$. Thus $Q_i \subseteq \mathrm{Ad}\, K(Q_i \cap h)$. Therefore $Q_1 = \mathrm{Ad}\, K(Q_1 \cap h) = \mathrm{Ad}\, K(Q_2 \cap h) = Q_2$.

(ii) $\Rightarrow$ (iii). Enough to show that for any $\mathrm{Ad}\,(K)$-invariant convex cone $Q \subseteq k$, we have $Q \cap h = p_h(Q)$. Clearly $Q \cap h \subseteq p_h(Q)$. If $X \in Q$, then for some $k \in K$, $H = \mathrm{Ad}\, k\, X \in Q \cap h$. Hence $p_h(X) = p_h(\mathrm{Ad}\, k^{-1} H)$. By Kostant's result $p_h(\mathrm{Ad}\, k^{-1} H) \in Q \cap h$.

We shall prove an analogous result for $\mathrm{Ad}\,(G)$-invariant cones in g.

Before doing this, we shall describe the cone $C_{\max}(g) \cap h$. Let $C^0_{\max}(g)$ (resp. $C^0_{\min}(g)$) denote the interior of $C_{\max}(g)$ (resp. $C_{\min}(g)$).

Proposition 4 : $C^0_{\max}(g) \cap h = \{H \in h : -i\alpha(H) > 0 \text{ for } \alpha \in P_n\}$.

Proof: We remark that since α is purely imaginary on h, $-i\alpha(H) \in R$ for any $H \in h$, and hence the condition above is meaningful.

To prove the proposition, we make use of the following *Theorem* (Vinberg). Let $X \in C^0_{\max}(g)$. Then X is conjugate to an element in $k^+ = \{Y \in k : J^{-1} \text{ ad } Y|_p$ is positive definite$\}$. Here p is endowed with the positive-definite metric $+B(,)$.

To go back to the proof of the proposition, let $H \in h$. Then $H \in C^0_{\max}(g) \cap h$ if and only if J^{-1} ad $H|_p > 0$. Now with the earlier notation, we remark that $\{X_\alpha, Y_\alpha : \alpha \in P_n\}$ is a basis for p and we can easily see that J^{-1}. ad H on p with respect to the above basis can be represented by the matrix

$$\begin{pmatrix} \frac{1}{i}\alpha_1(H) & 0 & & \\ 0 & \frac{1}{i}\alpha_1(H) & & \\ & & \frac{1}{i}\alpha_2(H) & 0 \\ & & 0 & \frac{1}{i}\alpha_2(H) \text{ etc.} \end{pmatrix},$$

where α_j runs through P_n and $i = \sqrt{-1}$. Hence J^{-1} ad H is positive-definite on p if and only if $1/i\,\alpha(H) > 0$, that is, if and only if $-i\alpha(H) > 0$ for all $\alpha \in P_n$.

Remark: Let C denote the closure of $C^0_{\max}(g) \cap h$ in h. That is $C = \{H \in h : -i\alpha(H) \geqslant 0$ for all $\alpha \in P_n\}$. Then it is easily seen that $C = C_{\max}(g) \cap h$.

Proposition 5: We have $p_h(C_{\max}(g)) = C_{\max}(g) \cap h = C$.

Proof: Since $C = C_{\max}(g) \cap h \subseteq h$, we have $C \subseteq p_h(C_{\max}(g))$. Thus we need only to show that $p_h(C_{\max}(g)) \subseteq C$.

Let $X \in C_{\max}(g)$. Let $X = A + B$ be the Cartan decomposition of X. Then for some $k \in K$, we have $\text{Ad}\,kA = H \in h$. Let us write $\text{Ad}\,kX = H + Y$, with $Y \in p$. Since $X \in C_{\max}(g)$, which is $\text{Ad}(G)$-invariant, we have $\text{Ad}\,k \cdot X \in C_{\max}(g)$. By the definition of $C_{\max}(g)$, we have

$$-B(\text{Ad}\,k(X), \text{Ad}\exp z \cdot (k_0)) \geqslant 0 \text{ for any } z \in p. \quad (*)$$

We now make a special choice of $z \in p$. We take $z = tX_\alpha$ for $\alpha \in P_n$ and $t \in R$. Then the above condition reads:

$$-B(\text{Ad}\exp(-t\,X_\alpha)(H), k_0) - B(Y, \text{Ad}\exp X_\alpha(k_0)) \geqslant 0.$$

Since k and p are orthogonal to each other with respect to $B(,)$ and in the first term above the vector $k_0 \in k$ to compute $-B(\text{Ad}\exp(-tX_\alpha)(H), k_0))$, it is enough to calculate the k-component in the Cartan decomposition of the vector $\text{Ad}\exp(-tX_\alpha)(H)$. Since $X \in p$, $H \in k$, $[k, p] \subseteq p$, and since $\text{Ad}\exp(-tX_\alpha)(H)$ is

$$(\exp \text{Ad}(-tX_\alpha))(H) = \sum_{0}^{\infty} \frac{(\text{ad}\,X_\alpha)^n(H)}{n!},$$

the k-component is given by

$$\sum_0^\infty \frac{(\operatorname{ad} X_\alpha)^{2n}(H)}{(2n)!}.$$

This is easily calculated, by observing

$$[H, X_\alpha] = \frac{1}{i}\alpha(H)\cdot Y_\alpha \text{ and } [H, Y_\alpha] = \frac{1}{i}\alpha(H)\cdot X_\alpha.$$

to be equal to

$$H + \left(\sum \frac{s^{2n}}{(2n)!}\cdot\frac{\alpha(H)}{2}\cdot 2^{2n}\right) H_\alpha,$$

$$= H + \frac{\alpha(H)}{2}(\cosh 2s - 1)\, H_\alpha, \text{ where } s = -t,$$

$$= H + \alpha(H) \sin h^2\, tH_\alpha.$$

A similar reasoning tells us that it is enough to calculate $(\operatorname{ad} X)^{2n+1}(k_0)$. Since $[X_\alpha, k_0] = -JX_\alpha$ and

$$(\operatorname{ad} X_\alpha)^3(k_0) = [X_\alpha [X_\alpha, -JX_\alpha]] = [X_\alpha [X_\alpha Y_\alpha]],$$

$$= 2[X_\alpha, iH_\alpha] = -2[iH_\alpha, X_\alpha]$$

$$= -2\cdot 2Y_\alpha = -2^2\cdot Y_\alpha, \text{ etc.}$$

We see that the p-component of Ad exp $tX_\alpha(k_0)$ is

$$-\frac{1}{2}\left(\sum \frac{t^{2n+1}}{(2n+1)!}\right) Y_\alpha = -\frac{1}{2}\sinh 2t \cdot Y_\alpha.$$

Thus (*) can be rewritten as

$$-B(H, k_0) + i\alpha(H)\sinh^2 t\, B(iH_\alpha, k_0)\ \tfrac{1}{2}\sinh 2t\ B(Y, Y_\alpha) \geqslant 0.$$

Dividing throughout by $\sinh^2 t$ and using the fact that $\frac{1}{2}\sinh 2t = \sinh t.\cosh t$, we rewrite this as

$$-\frac{1}{\sinh^2 t}B(H, k_0) + i\alpha(H)\,B(iH_\alpha, k_0) \geqslant -\coth t\, B(Y, Y_\alpha).$$

Letting $t \to \pm\infty$, we get

$$-i\alpha(H)(-B(iH_\alpha, k_0)) \geqslant \mp B(Y, Y_\alpha).$$

Since $-B(iH_\alpha, k_0) > 0$, we see that $-i\alpha(H) \geqslant 0$. This completes the proof of the proposition.

Let $T_\alpha = iH_\alpha \in h$. Since $-B(,)$ restricted to h is positive definite, we identify h' with h *via* $-B(,)$, under this identification $t_\alpha = T_\alpha/[+B(\alpha, \alpha)]$ corresponds to $-i\alpha$. Now from lemma 1 we see that the dual C^* of $C \subseteq h$ can be described as follows:

$$C^* = \{\sum_{\alpha\in Pn} a_\alpha t_\alpha\ ;\ a_\alpha \geqslant 0\} = \{\sum_{\alpha\in Pn} b_\alpha T_\alpha : b_\alpha \geqslant 0\}.$$

We now want to prove $p_h(C_{\min}(g)) = C^*$. For this we need to set up some notation.

We say that two roots α and α' in P_n are strongly orthogonal if neither $\alpha + \alpha$ nor $\alpha - \alpha'$ is in $\triangle$. A subset $S \subseteq P_n$ is said to be strongly orthogonal if and only if any two roots in S are strongly orthogonal. It is well-known that there exists a (maximal) strongly orthogonal subset $S \subseteq P_n$ such that the space

$$a_0 = \sum_{\gamma \in S} R(Z_\gamma + Z_{-\gamma})$$

is a maximal abelian subspace of p [3; see also 5, Prop. VIII. 7.4 Cor. 7.6].

Lemma 6 : Let $Z = \sum_{\gamma \in S} t_\gamma (Z_\gamma + Z_{-\gamma}) \in a$, with $t_\gamma \in R$.

Let $H \in h$. Then we have

$$(\mathrm{Ad}\ \exp Z)(H)$$
$$= H + \Sigma_S\, \gamma(H) \sinh t_\gamma \cosh t_\gamma (Z_{-\gamma} - Z_\gamma) + \Sigma_S \sin h^2 t_\gamma (\gamma(H) \cdot H_\gamma).$$

Proof : For Z as above, we can write $\exp Z = \exp Y \exp A \exp X$ where $Y = \Sigma_S$ $\tan\ t_\gamma Z_{-\gamma}$, $X = \Sigma_S \tan\ t_\gamma Z_\gamma$
and $A = \Sigma_S \log(\cosh t_\gamma)\ [Z_\gamma\ Z_{-\gamma}]$ [5; Lemma VIII. 7.11].

We then easily see that

$$\mathrm{Ad}\ \exp X(H) = H - \Sigma\, \gamma(H) \cdot \tanh t_\gamma Z_\gamma$$
$$(\mathrm{Ad}\ \exp A \circ \mathrm{Ad}\ \exp X)(H) = H - \Sigma\, \gamma(H) \tanh t_\gamma \cosh^2 t_\gamma Z_\gamma,$$
$$= H - \Sigma\, \gamma(H) \sinh t_\gamma \cosh t_\gamma Z_\gamma,$$
$$(\mathrm{Ad}\ \exp Y \mathrm{Ad}\ \exp A\ \mathrm{Ad}\ \exp X)(H)$$
$$= H + \Sigma\, \gamma(H) \tanh t_\gamma Z_{-\gamma} - \Sigma\, \gamma(H) \sinh t_\gamma \cosh t_\gamma (Z_\gamma - \tanh t_\gamma H_\gamma$$
$$- \tanh^2 t_\gamma Z_{-\gamma}),$$
$$= H + \Sigma\, \gamma(H) \sinh t_\gamma \cosh t_\gamma (Z_{-\gamma} - Z_\gamma) + \Sigma \sinh^2 t_\gamma (\gamma(H) \cdot H_\gamma).$$

This completes the proof of the lemma.

We let p_k denote the projection of g onto k with respect to the Killing form $-B(,)$. We note that for $k \in K$, and $X \in g$, we have $(p_k \circ \mathrm{Ad} k)(X) = \mathrm{Ad}\, k\,(p_k(X))$. Also, we have $\gamma(k_0) = i$ for $\gamma \in P_n$.

Corollary 7 : With the notation of the lemma 6, we have

$$p_k(\mathrm{Ad}\ \exp Z(H)) = H + \sinh^2 t_\gamma (\gamma(H) \cdot H_\gamma).$$
$$= p_h(\mathrm{Ad}\ \exp Z(H)).$$

In particular for $H = k_0$, we have

$$p_k(\mathrm{Ad}\ \exp Z(k_0)) = k_0 + \Sigma \sinh^2 t_\gamma (iH_\gamma)$$
$$= k_0 + \Sigma \sinh^2 t_\gamma T_\gamma = p_h(\mathrm{Ad} \exp Z(k_0)).$$

Moreover for any $Z \in a$, we have

$$p_k(\mathrm{Ad}\ \exp Z(k_0)) \in C^*.$$

Proof : The first statement follows from the above lemma, since $\gamma(H)$ is purely imaginary, $\gamma(H)(Z_{-\gamma} - Z_\gamma) \in p$. The last statement follows from the expression for $p_h(\mathrm{Ad}\ \exp Z(k_0))$ and the description of C^* given before the statement of lemma 6,

Proposition 8 : $p_h(C_{\min}(g)) = C^*$.

Proof : Since $C_{\min}(g)$ = the closed convex cone generated by Ad $(\exp X)(k_0)$ for $X \in p$, it is enough to show that p_h (Ad exp X) $(k_0) \in C^*$. By the corollary 7, if $X \in a$, then we are through. We now want to exploit the fact that

$$p = \bigcup_{k \in K} \mathrm{Ad}\, k(a) \text{ to prove the general case.}$$

We first prove the following *claim* :

$$p_k(C_{\min}(g)) = \mathrm{Ad}\, K(C^*).$$

For this it is enough to check that p_k (Ad exp $X \cdot k_0$) = Ad $K(C^*)$ for $x \in p$. $X \in p$ can be written as Ad k (H) for some $k \in K$ and $H \in a$. Thus we have only to show that

$$p_k(\mathrm{Ad}(\exp(\mathrm{Ad}\; k(H)))\, k_0) \in \mathrm{Ad}\, K(C^*) \text{ for all } k \in K \text{ and } H \in a.$$

Now the element under inspection is p_k (Ad k Ad exp $H \cdot k_0$). Since $p_k \circ \mathrm{Ad}\, k = \mathrm{Ad}\, k \circ p_k$, we have to check whether p_k (Ad k Ad exp Hk_0) = Ad $k\, p_k$ (Ad exp $H \cdot k_0$) $\in$ Ad $K(C^*)$. By the corollary 7, we have

$$p_k(\mathrm{Ad}\exp Hk_0) = p_h(\mathrm{Ad}\exp Hk_0) \in C^* \text{ and hence}$$

$$\mathrm{Ad}\, k\, p_k(\mathrm{Ad}\exp Hk_0) \in \mathrm{Ad}\; k(C^*) \subseteq \mathrm{Ad}\, K(C^*).$$

Thus we have proved $p_k(C_{\min}(g)) \subseteq \mathrm{Ad}\, K(C^*)$. To prove the other inclusion, it is enough to show that $C^* \subseteq C_{\min}(g)$. We note that C^* is a W_K-invariant convex cone in h. Let $H \in C^*$. Since $C_{\min}(g) = (C_{\max}(g))^*$, the dual of $C_{\min}(g)$ with respect to $-B(,)$, we have only to check that $-B(H, X) \geqslant 0$ for any $X \in C_{\max}(g)$. Since

$$H \in C^* \subseteq h, \; -B(H, X) = -B(H, p_h(X)).$$

Since $p_h(X) \in C$ by proposition 4 and C^* is the dual of C, we deduce that $-B(H, p_h X) \geqslant 0$. This completes the proof of the claim.

We now see that

$$\begin{aligned} p_h(C_{\min}(g)) &= p_h(p_k(C_{\min}(g))), \\ &= p_h(\mathrm{Ad}\, K(C^*)) \text{ (by the claim above)}, \\ &= C^* \text{ (by corollary 3)}. \end{aligned}$$

This completes the proof of the proposition.

Remark : Note that in view of corollary 3, we have shown in the course of the above proof, that

$$p_h(C_{\min}(g)) = C^* = C_{\min}(g) \cap h.$$

In fact, for any invariant convex cone $Q \subseteq g$, the proof of the claim can be adapted to show $p_h(Q) = Q \cap h$. We treated here the special case when $Q = C_{\min}(g)$ to bring out the salient features of the proof. The general case is formulated in a more precise way as theorem 11.

Corollary 10 : C^* is the closed convex cone generated by

$$\{T_{w\gamma} : \gamma \in S,\; w \in W_K\}.$$

Proof: This follows from corollary 7 and proposition 8. Note that k_0 being the average of T_w's already lies in the cone generated by

$$\{T_{w\gamma} : \gamma \in S,\ w \in W_K\}.$$

We can now describe all invariant convex cones in g. We first observe that if Σ is any subset of h we denote by $C(\Sigma, h)$ the convex cone generated by $\{w \cdot X ;\ w \in W_K,\ X \in \Sigma\}$ in h. Note in every case for any X the element $\sum_w w \cdot X$ is an element of h that is left fixed by every element of W_K. Hence it has to be a scalar multiple of k_0. Now, since $C(\Sigma, h)$ is a W_K-invariant convex cone in h, by lemma 2, Ad $K(C(\Sigma, h))$ is an Ad(K)-invariant cone in k such that $p_h(\text{Ad}\, K(C(\Sigma, h))) = \text{Ad}\ K(C(\Sigma, h)) = C(\Sigma, h)$. Let $Q(\Sigma)$ denote the closed convex Ad(G)-invariant cone in g generated by $\Sigma \subseteq h$. We can now put together all the pieces to state and prove our main result.

Theorem 11: Let C_0 be a W_K-invariant convex cone in h such that $C^* \subseteq C_0 \subseteq C$. Let $Q(C_0)$ denote the Ad(G)-invariant convex cone in g generated by C_0. We then have

$$p_h(Q(C_0)) = Q \cap h = C_0.$$

Conversely given an Ad(G)-invariant convex cone Q in g, such that $k_0 \in Q$, Q is generated by $Q \cap h = p_h(Q)$.

Proof: The proof is exactly similar to that of proposition 8 and uses lemmas 6 and 2 (or rather corollary 3).

To prove the first statement, since obviously $C_0 \subseteq p_h(Q(C_0))$, it is enough to prove that $p_h(\text{Ad}\, g(H)) \in C_0$ for $H \in C_0$, $g \in G$. This follows from the following claim:

Claim 1: If $H \in C^\circ$, the interior of C, and $g_0 \in G$, then $p_h(\text{Ad}\ g_0(H_0)) \in$ convex-hull of $\{W_K \cdot H_0\} + C^*$.

To prove this claim we need a lemma.

Lemma: Let $X \in C^0_{\max}(g)$. Then either $X \in k^+$ or there exists an element Y in the orbit of X such that

(i) $p_h(X) \in$ convex-hull of $\{W_K \cdot p_h(Y)\} + C^*$,

(ii) $|p_p(Y)| < |p_p(X)|$

where p_p denotes the projection on the p-component of $g = k \oplus p$ and $|p_p(X)|$ is the norm of $p_p(X)$ in the Killing norm.

Proof: Let $X \notin k^+$. By conjugating by an element of Ad K, we may assume that

$$X = H + \sum_{\alpha \in P_n} (t_\alpha X_\alpha + t_{-\alpha} Y_\alpha), \text{ with } H \in C^0.$$

$$\text{By taking } Z = i \sum \left(\frac{t-\alpha}{\alpha(H)}\ X_\alpha - \frac{t_\alpha}{\alpha(H)}\ Y_\alpha \right)$$

we see that

$$[Z, X] = - \sum_{\alpha \in P_n} (t_\alpha X_\alpha + t_{-\alpha} Y_\alpha) + S$$
$$+ (i)^2 \sum_{\alpha \in P_n} \left[\frac{t_{-\alpha}^2}{-i\alpha(H)} + \frac{t_\alpha^2}{-i\alpha(H)} \right] iH_\alpha, \text{ where } S \in h^\perp \cap k$$

Put

$$V = \sum_{\alpha \in P_n} \left[\frac{t_{-\alpha}^2}{-i\alpha(H)} + \frac{t_\alpha^2}{-i\alpha(H)} \right] iH_\alpha. \text{ Notice that } V \in C^*.$$

Hence for $t \in R$, we can write

$$\operatorname{Ad} \exp tZ(X) = H + (1 - t) \sum_{\alpha \in P_n} (t_\alpha X_\alpha + t_{-\alpha} Y_\alpha) - V + tS + \psi(t^2),$$

where $\psi(t^2)$ is the sum of terms involving t^2. Hence for $t > 0$ and sufficently small we have

$$H - p_h(\operatorname{Ad} \exp tZ(X) = p_h(X) - p_h(\operatorname{Ad} \exp tZ(X)) \in C^*.$$

We take $Y = \operatorname{Ad} \exp tZ(X)$ with t as above. Then $|p_{\mathfrak{p}}(Y)|$ is dominated by $(1-t)|p_{\mathfrak{p}}(X)|$ and hence $|p_h(Y)| < |p_{\mathfrak{p}}(X)|$.

The general case follows by the convexity theorem of Kostant: If X_0 is the given element and if $\operatorname{Ad} k(X_0) = X$ is as above, then we have

$$\begin{aligned} p_h(X_0) &= p_h(\operatorname{Ad} k^{-1} \operatorname{Ad} k(X_0)) = p_h(p_k(\operatorname{Ad} k^{-1} \operatorname{Ad} k(X_0))), \\ &= p_h(\operatorname{Ad} k^{-1} p_k(\operatorname{Ad} k(X_0))), \\ &= p_h(\operatorname{Ad} k^{-1} p_h(\operatorname{Ad} k(X_0))), \text{ since } p_h(X) = p_k(X). \end{aligned}$$

Thus the lemma is completely proved.

Proof of the claim: Let $X = \operatorname{Ad} g_0(H_0)$ and $p_h(X) = H \in C^\circ$. Consider the orbit $O_X = \{\operatorname{Ad} g X : g \in G\}$. Since X is semi-simple, O_X is closed by a well-known result of Borel–Harish–Chandra [10]. In this orbit consider the set

$$O_X^+ = \{Y \in O_X : H \in \text{convex-hull of } \{W_K \cdot p_h(Y)\} + C^*$$
$$\text{and } |p_{\mathfrak{p}}(Y)| < |p_{\mathfrak{p}}(X)|\}.$$

We claim that O_X^+ is compact. Since it is obviously closed, only its boundedness to be proved, which we relegate to a lemma below. We now consider the function $Y \mapsto |p_{\mathfrak{p}}(Y)|$ on O_X^+ to $\mathcal{R}$. Since O_X^+ is compact, it attains its minimum say at Y. We contend that $p_{\mathfrak{p}}(Y) = 0$, since otherwise, the lemma above will give us an element Y' such that $Y' \in O_X^+$ and $|p_{\mathfrak{p}}(Y')| < |p_{\mathfrak{p}}(Y)|$. Hence $p_{\mathfrak{p}}(Y) = 0$. Thus $Y \in k^+$. If we take $H_1 \in C^0$ to be such that $\operatorname{Ad} k\, Y = H_1$, it then follows from what we have proved that H_1 and H_0 lie in $C^0 \subseteq h$, are conjugate and are such that

$$H_0 \in \text{convex-hull of } \{W_K \cdot H_1\} + C^*.$$

But then convex-hull of $\{W_K \cdot H_1\}$ = convex-hull of $\{W_K \cdot H_0\}$. Hence the claim is completely proved modulos the assertion that O_X^+ is compact.

That O_X^+ is bounded follows from the following:

Lemma: Let $H \in k$ and C a constant > 0 and $g \in G$. If $|p_{\mathfrak{p}}(\operatorname{Ad} g(H))|^2 < C$, then $|\operatorname{Ad} g(H)|^2 < 2C + |H|^2$ where $|X|$ is with respect to $-B(X, \theta X)^{1/2}$,

Proof: Put $g = k\exp(X), k \in K, X \in p$.

Then $|p_p(\mathrm{Ad}\, g(H))| = |p_p \mathrm{Ad}\exp(X)(H)|$. Since $X \in p$, we can choose an orthonormal basis $e_1, \ldots, e_r$ of k and $f_1, \ldots, f_s$ of p such that

$$\left.\begin{aligned} \mathrm{Ad}\, X(e_i) &= \lambda_i f_i \\ \mathrm{Ad}\, X(f_i) &= \lambda_i e_i \end{aligned}\right\} \; 1 \leqslant i \leqslant \min(r, s).$$

All other basis elements, if any, go to zero under $\mathrm{Ad}\, X$. Let $H = \Sigma a_i e_i$. For simplicity we assume $r = s$. Hence we have

$$\mathrm{Ad}\exp X(H) = \Sigma a_i (\cosh \lambda_i e_i + \sinh \lambda_i f_i)$$

and therefore

$$\begin{aligned} |\mathrm{Ad}\exp(X)(H)|^2 &= \Sigma a_i^2 (\cosh^2 \lambda_i + \sinh^2 \lambda_i), \\ &= |H|^2 + 2\Sigma a_i^2 \sinh^2 \lambda_i. \\ &< |H|^2 + 2C, \end{aligned}$$

since

$$|p_p(\mathrm{Ad}\exp X(H))|^2 = \Sigma a_i^2 \sinh^2 \lambda_i.$$

This completes the proof of the lemma and hence the claim.

To prove the second part of the theorem let Q be the given invariant convex cone containing k_0. Let $X \in Q^0$, the interior of Q. Since $Q \subseteq C_{\max}(g)$, by Vinberg's theorem [9; theorem 5] there exists $g \in G$ such that $\mathrm{Ad}\, g X \in k^+$. But then by a suitable $\mathrm{Ad}(k)$-conjugation we conclude that $\mathrm{Ad}\, k g(X) \in h$. Since Q is $\mathrm{Ad}(G)$-invariant, we have $\mathrm{Ad}\, kg(X) \in Q \cap h$. This means that $\mathrm{Ad}\, kg(X)$ lies in the $\mathrm{Ad}(G)$-invariant cone, generated by $Q \cap h$ and hence X itself lies in the invariant cone generated $Q \cap h$. By the first part of the theorem we have $p_h(X) \in Q \cap h$. This completes the proof of the second part.

The following corollary is immediate from theorem 11:

Corollary: Let g be such that $C_{\min}(g) \neq C_{\max}(g)$. (This is the same as saying $C^* \neq C$, in our notation.) Then there exists a continuum of invariant convex cones in g.

We note that this corollary answers Vinberg's problem 1 [9], in the negative.

We can also answer Vinberg's question 4 [9] in the affirmative:

Theorem 11 " describes " all the invariant convex cones in a simple Lie algebra in terms of the intersection with the compact Cartan subalgebra h. The second part of the problem 4 asks whether it is true $Q^* \cap h = (Q \cap h)^*$. The answer to this question is yes. This also follows trivially from theorem 11 as shown below:

Cones Q and Q^* are determined by $Q \cap h$ and $Q^* \cap h$ by theorem 11. Note that we have trivially $Q^* \cap h \subseteq (Q \cap h)^*$ since if $X \in Q^* \cap h$, $Y \in Q \cap h$, then we have $-B(X, Y) \geqslant 0$. Now let $H \in (Q \cap h)^*, X \in Q$. We need only show that $-B(H, X) \geqslant 0$. But then $-B(H, X) = -B(H, p_h(X))$. Now by theorem 11, $p_h(X) \in Q \cap h$ and so $-B(H, p_h(X)) \geqslant 0$ since $H \in (Q \cap h)^*$. This implies $-B(H, X) \geqslant 0$ for all $X \in Q$. That is to say that $H \in Q^*$. This gives the complete answer to Vinberg's problem 4. This we formulate as

Proposition 12 : For any Ad (G)-invariant convex cone $Q \subseteq g$, we have $(Q \cap h)^* = Q^* \cap h$. Hence Q is self-dual if and only if $Q \cap h$ is a self-dual cone in h.

Remark : Using our description of C and C^* in h, one can easily verify that $C = C^*$ when and only when $g = \mathrm{sp}(n, R)$. In this case, if h is identified with R^n, then $C = C^* =$ the 'quadrant' $\{x \in \mathcal{R}^n : x_i \geqslant 0\}$. Hence the only case for which $C_{\min}(g) = C_{\max}(g)$ is true occurs when $g = \mathrm{sp}(n, R)$. This is noticed by Vinberg [9], but his proof is not clear to us.

Remark : One may now ask : Are there self-dual cones in g? If they exist, are there only finitely many self-dual cones in g? The answer to the former is yes. This can be easily seen by our explicit description of C and C^*. All one has to observe is that one can squeeze a W_K-stable self-dual cone C_0 such that $C^* \subseteq C_0 \subseteq C$. The reader may easily verify this claim.

But the answer to the second question is No. To show this we consider the case of SO* (6) (which is locally SU (3, 1), but to identify h with R^3, SO* (6) realization is easier to handle). With the standard notation of the root system $D_3 (\sim A_3)$ we may take as the unique noncompact simple root $e_2 + e_3$ (The simple roots are $e_2 + e_3$, $e_1 - e_2$, $e_2 - e_3$). Then we have

$$P_n = \{e_1 + e_2, e_2 + e_3, e_1 + e_3\}.$$

Then identifying h with R^3 with the standard inner product, we see that

$$C = (x, y, z) \in R^3 : \ x + y \geqslant 0,\ y + z \geqslant 0 \text{ and } x + z \geqslant 0$$

and

$$C^* = \{(x, y, z) \in R^3 : \ y + z \geqslant x,\ z + x \geqslant y \text{ and } x + y \geqslant z\}.$$

The Weyl group W_K can be identified with S_3 which acts on R^3 in the usual manner. In the appendix we construct for any $n \geqslant 1$, a family of cones C_r^n such that one member of this family is self-dual. For $n \neq m$, $C_r^n \neq C_s^m$ for any r, s. This construction shows that SO* (6) has an infinite number of self-dual invariant convex cones.

Remark : We want to record some observations we have made on the cones in h. In the followin by a cone in h or g we mean its interior. It is easily seen that if a W_K-stable convex cone $C \subseteq h$ is homogeneous in h (for definition, (see [8]), then $Q(C)$, the Ad (G)-invariant cone generated by C in g is also homogeneous. If we assume further that C is self-dual too, then $Q(C)$ is a homogeneous self-dual cone. Notice that $Q(C)$ is *a priori* self-dual with respect to $-B$. But as is easily seen $Q(C)$ is self-dual with respect to $-B_\theta$ where $-B_\theta(X, Y) = -B(X, \theta Y)$, i.e., $Q(C)$ is a self-dual homogeneous cone in a 'Euclidean' space $(g, -B_\theta)$. Then by a known result [1, 7], $Q(C)$ is a Riemannian symmetric space with respect to the canonical Riemannian structure on it (for the definition of the canonical Riemannian structure see [7, 8]). It would be interesting to know which of the self-dual W_K-invariant cones C are homogeneous and find all the Riemannian symmetric spaces associated to those cones $Q(C)$ in g. For example, in the case when $G = \mathrm{sp}(n, \mathcal{R})$ we have seen that $C_{\min}(g) = C_{\max}(g) = Q$. If one interprets $\mathrm{sp}(n, \mathcal{R})$ as quadratic forms on $\mathcal{R}^{2n}$, then Q may be identified with those form that are positive definite with respect to the sympletic form [which defines $\mathrm{sp}(n, \mathcal{R})$]. Thus in this case, the corresponding symmetric space is $\mathcal{R}^+ \times \mathrm{SL}(2n, \mathcal{R})_{SO(2n)}$.

Appendix

In this appendix we discuss some very special type of convex cones in a finite dimensional Euclidean space and describe their duals. Then we shall further specialize to discuss a class of cones, their duals and the conditions for their self-duality.

Let V denote a finite dimensional vector space over $\mathscr{R}$ and V' its real dual. By a convex cone in V we mean a closed convex cone which has interior and which contains no straight line entirely. We assume that V is endowed with an inner product $\langle\, ,\rangle$. If C is a cone in V, its dual, denoted by C^*, is defined as follows:

$$C^* = \{v' \in V' : v'(v) \geqslant 0 \text{ for all } v \text{ in } C\}.$$

Then it is well-known that $(C^*)^* = C$ and that for two cones $C_1 \subseteq C_2$, we have $C_1^* \supseteq C_2^*$. Using the inner product $\langle\, ,\rangle$ we can and do identify V' with V.

Lemma1 : Let $\{v_1, \ldots, v_N\}$ be a set of vectors in V such that $C = \{v \in V : \langle v, v_i \rangle \geqslant 0 \text{ for } 1 \leqslant i \leqslant N\}$ is a cone in V. Then we have

$$C^* = \{u \in V : u = \sum_{1 \leqslant i \leqslant N} a_i v_i,\ a_i \geqslant 0.$$

Proof : It is easily seen that

$$C' = \{u \in V : u = \sum_{1 \leqslant i \leqslant N} a_i v_i,\ a_i \geqslant 0\}$$

is indeed a cone. Let C^* be the dual of C. Then it is clear that $C' \supseteq C^*$. Taking duals, we see that $C \subset (C')^*$. Hence to show that $C' \supseteq C^*$, it is enough to show that $(C')^* \subseteq C$. Now if $u \in V$ is such that $\langle u, \Sigma\, a_i v_i \rangle \geqslant 0$ for any $\Sigma\, a_i v_i \in C'$, then it follows that $\langle u, v_i \rangle \geqslant 0$ for all i. That is, $u \in C$.

We now discuss some very special type of cone in $\mathscr{R}^3$ with the usual inner product.

We call a cone C a regular polygonal right cone (in short RPRC) if there exists a plane P not passing through the origin O such that

(i) $P \cap C$ is a regular polygon called the base,

(ii) the line joining the centroid G of $P \cap C$ to O is orthogonal to P. We shall call this line the axis of the cone (figure 1).

We shall describe the duals of such cones. For simplicity we shall assume that P is the plane $z = 1$ in $\mathscr{R}^3 = \{(x, y, z) :\ x, y, z \in R\}$ and it is the plane of the paper.

Suppose now that $P \cap C$ is a regular $(n+1)$-gon, its centroid is the point $(0, 0, 1)$ and that one of its vertices v_0 lies on the line $y = 0, z = 1$. Let r be the distance of the vertices $\{v_k\}_0^n$ from G. (Recall that the vertices are equidistant from G). With these normalizations, which do not result in any loss of generality, we can write

$$v_k = \left(r \cos \frac{2\pi k}{n+1},\ \underline{r} \sin \frac{2\pi k}{n+1},\ 1\right),\quad 0 \leqslant k \leqslant n.$$

Let $C_r = \{(\Sigma\, a_k v_k,\ a_k \geqslant 0\}$. Then C is an RPRC. By Lemma 1, we have $C_r^* = \{(x, y, z) \in \mathscr{R}^3 : \langle (x, y, z), v_k \rangle \geqslant 0;\ 0 \leqslant k \leqslant n]$.

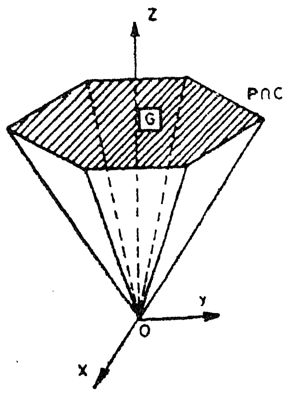


Figure 1

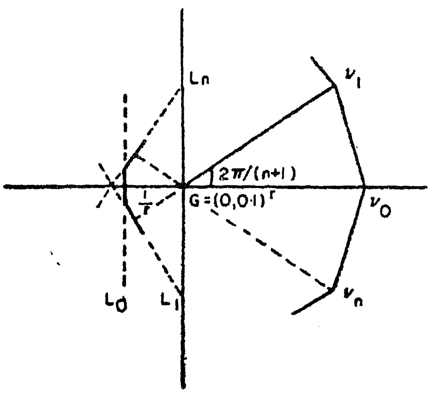


Figure 2

Thus the bounding hyperplanes of C^* are

$$r\cos\frac{2\pi k}{n+1}x + r\sin\frac{2\pi k}{n+1}y + z = 0,\ 0 \leqslant k \leqslant n.$$

Intersection of these planes with the plane $P : \{z = 1\}$ gives rise to an $(n + 1)$-gon whose sides are given by the equations

$$L_k : r\cos\frac{2\pi k}{n+1}\cdot x + r\sin\frac{2\pi k}{n+1}y + 1 = 0,\ 0 \leqslant k \leqslant n.$$

L_0 and L_1 are depicted in the diagram (figure 2). Note also that the equation for L_0 is $rx + 1 = 0$ or $x = -1/r$ and that the slope of L_k is $-\cot 2\pi k/n + 1$ which is independent of r.

We now discuss two cases, of which the first one is what is needed in the main part of the paper.

Case 1 : $n+1$ is odd, say $2m+1$. In this case the sides of the base of C^* are parallel to that of the base of C. Hence the dual C_r^* of any $C_r \in \{C_r\}$ again lies in the family, *i.e.*, $C_r^* = C_s$ for some s (figure 3). In fact, s can be computed explicitly :

$$\frac{s}{r} = \left(\frac{1}{n}\right)/r \cos(\pi/2m+1),$$

or

$$s = \frac{1}{r}(\sec\ (\pi/2m+1)).$$

That is,

$$C_r^* = C_{1/r}\,(\sec \pi/2m+1).$$

In particular if $r_0 = (\sec \pi/2m+1)^{1/2}$, then

$$C_{r_0}^* = C_{r_0},\ \text{i.e., } C_{r_0} \text{ is self-dual.}$$

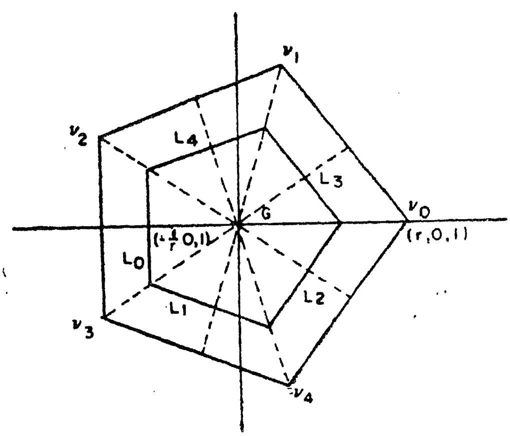
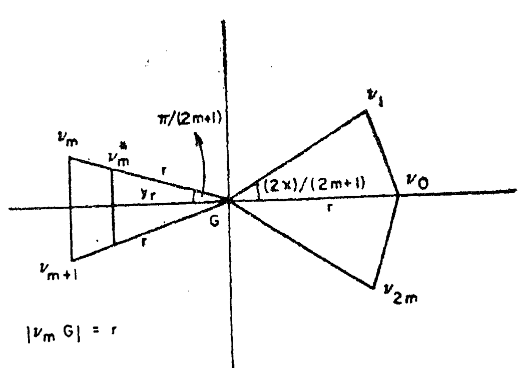


Figure 3



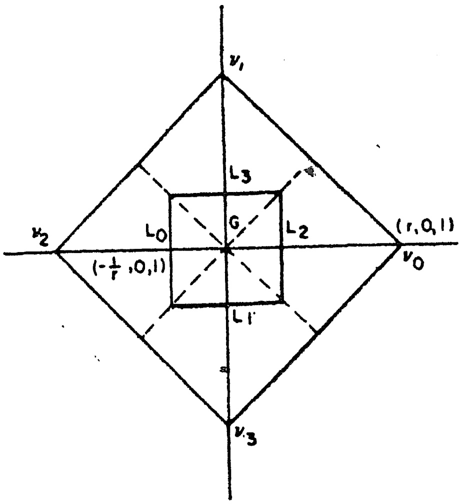


Figure 5

Case 2 : $n+1$ is even, say, equal to $2m$. In this case, it can be checked that the dual C_r^* of C_r has as its base a regular $2m$-gon whose sides are not parallel to the bases of cones in C_r (figure 4). In particular the families $\{C_r^*\}$ and $\{C_r\}$ are distinct and hence no C_r can be self-dual.

We now let S_3 the symmetric group on $\{1, 2, 3\}$ act on $\mathscr{R}^3$ in the usual way, *i.e.*, if $(x_1, x_2, x_3) \in R^3$ then for $S_3\, \sigma(x_1,\ x_2\, x_3) = (x_{\sigma(1)}\ x_{\sigma(2)}\, x_{\sigma(3)})$. Then it is clear that the cones $C_r^{3(2n+1)}$ constructed above are S_3-stable (the superscript denotes the number of sides of the base polygon).

References

[1] Dorfmeister J and Koecher M Regulare Kegel 1979 *J. Deutsch Math. Verein* **81** 109–151

[2] Duistermaat J J, Kolk J A C and Varadarajan V S 1981 Preprint No. 200, University Utrecht

[3] Harish-Chandra 1956 *Am. J. Math.* **78** 564–628

[4] Heckman G J 1980 Projection of orbits and asymptotic behaviour of multiplicities for compact Lie groups. Thesis Rijksuniversiteit Leiden

[5] Helgason S 1978 *Differential geometry, Lie groups and symmetric spaces* (New York ; Academic Press)

[6] Kostant B 1973 *Ann. Sci. Ec. Norm. Sup.* **6** 413–455

[7] Rothaus O S 1960 *Ann. Math.* **83** 358–376

[8] Vinberg E B 1963 *Trans. Moscow Math. Soc.* **12** 340–403

[9] Vinberg E B 1980 *Invariant convex cones and orderings in Lie groups, functional analysis and its applications* *Eng. Trans.* Vol. 14, 1–10

[10] Warner G 1972 *Harmonic analysis on semisimple Lie groups* Vol. I (Berlin : Springer Verlag)

Proc. Indian Acad. Sci. (Math. Sci.), Vol. 91, Number 3, November 1982, pp. 183–193.

Minimum error solutions of boundary layer equations

NOOR AFZAL
Department of Mechanical Engineering, Aligarh Muslim University, Aligarh 202 001, India

MS received 17 September 1980

Abstract. The minimum error solutions of boundary layer equations in the least square sense have been studied by employing the Euler–Lagrange equations. To test the method a class of problems, *i.e.*, boundary layer on a flat plate, Hiemenz flow, boundary layer on a moving sheet and boundary layer in non-Newtonian fluids have been studied. The comparison of the results with approximate methods, like Karman–Pohlhuasen, local potential and other variational methods, shows that the present predictions are invariably better.

Keywords. Minimum error solutions ; boundary layer equations ; Euler–Lagrange equation.

1. Introduction

The approximate solutions of boundary layer equations based on classical Karman's integral have been extensively studied in the literature. The exposition of current state of knowledge may be found in the monographs of Rosenhead [10] and Schlichting [14].

Motivated by the success of variational methods in various branches of physics and engineering many workers have attempted the approximate solutions of boundary layer equations *via* variational formulations. Schechter [12, 13], Doty and Blick [1] and Venkateshwarlu [16] using the local potential theory [2, 9, 4, 6], have studied boundary layer flows and the results were better than the Karman-Pohlhausen method.

The local potential methods are based on a general evolution criteria where the dynamical system is in a stationary state when the functional (called local potential) is minimum. The minimisation of the functional through Euler–Lagrange equations along with certain auxiliary conditions leads to the conservation equations. The local potential theory employs two kinds of variables, the thermodynamic variables (temperature T, velocity u, pressure p, etc.) that change during the variational process and the variables of same type (T^0, u^0, p^0, etc.) evaluated at the stationary state that are not subjected to variation. This dual character of

the variables must be maintained until the variational process is complete. After setting the auxiliary conditions $T = T^0$, $u = u^0$, $p = p^0$, etc., yields the correct conservation equations. The local potential variational formulation is nonclassical due to the dual character of variables when compared to the classical theory of the calculus of variations. The local potential method of Glansdorff and Prigogine [2] and Prigogine and Glansdorff [9] may be regarded as a generalised entropy production whose production rate at the stationary state is minimal. The method of Lambermont and Lebon [6] and Lebon and Lambermont [7] is based on an extension to Hamilton's principle and the central quantity is the Legendre transform of internal energy per unit volume.

In §2 it is shown that all the local potential methods are equivalent to a particular moment of the boundary layer equations and for a given velocity profile the different local potential formulations will lead to the same results. It is also possible to choose another moment function but then method becomes arbitrary.

Vujanovic and Djukic [17] have developed a different approach to construct a Lagrangian for the boundary layer equations. Their functional contains a parameter, with no physical interpretation, which has to be made to tend to zero after the variation is complete, to get the correct boundary layer equations. The method has been applied to study the flat plate boundary layer problem and for a particular velocity profile the results are better than Karman's integral method. For other velocity profile, however, the method fails to produce good results [16].

Hsu [5] employed the Galerkin method to study a class of two-dimensional boundary layer problems. It can be shown that the local potential or the Galerkin method for a given problem, will yield same results when the approximating functions are the members of a complete set. When the solutions are not based on complete set, the functional has to satisfy certain additional conditions in terms of unknown parameters to get the results equivalent to Galerkin's technique [8].

In all the methods mentioned above, one has to select a proper velocity profile with one or more unknown parameters. The unknown parameters are determined in different ways depending on the method used. As long as one has to start with an assumed form of velocity profile with some free variables, one might determine the parameters by minimizing the error in the least square sense.

The aim of the present work is to develop a general method to determine the minimum error solutions of boundary layer equations in the least square sense. In the present method the integral of the square of the residual of the boundary layer equations is minimized by employing the well-known Euler–Lagrange equations from the calculus of variations. To test the method various problems in viscous flows, *viz.*, boundary layer on flat plate, Hiemenz flow, boundary layer on the moving sheet, boundary layer in non-Newtonian fluids are studied and it is found that the method gives invarialby better results.

2. Boundary layer equations and local potential proposals

The laminar two-dimensional boundary layer flow of an incompressible fluid with constant properties is governed by the following equations:

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0, \tag{1}$$

$$u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} = U\frac{\partial U}{\partial x} + \frac{\partial \tau}{\rho \partial y}. \tag{2}$$

The appropriate boundary conditions are

$$y = 0, \quad u = v = 0, \tag{3a,b}$$

$$y \to \infty, \quad u \to U(x). \tag{4}$$

Here x is the distance measured along the body, y normal to it. u and v are the velocity components in x and y directions respectively. $U(x)$ is local free stream velocity, ν is the kinematic viscosity and τ is the local shear stress given by

$$\tau = \mu \frac{\partial u}{\partial y}. \tag{5}$$

For flow past a flat plate the various local potential formulations proposed by various authors are given below :

(a) *Glansdorff and Prigogine* [2]

$$I_{GP} = \int_0^l \int_0^\infty (- u^0 u^0 u_x - v^0 u^0 u_y + \nu/2\, u_y^2)\, \mathrm{d}y\, \mathrm{d}x + \int_0^\infty u^0 u^0 u \Big|_{x=l} \mathrm{d}y. \tag{6}$$

(b) *Lebon–Lambermont*

$$I_{LL} = \int_0^l \int_0^\infty \left[u^2 u_x^0 + uvu_y^0 + \frac{u^{0^2}}{2}(u_x + v_y) + \nu/2\, u_y^2\right] \mathrm{d}x\, \mathrm{d}y$$

$$- 1/2 \int_0^\infty u^3 \Big|_{x=l} \mathrm{d}y. \tag{7}$$

(c) *Hiroaka–Tanaka* [4]

$$I_{HT} = \int_0^l \int_0^\infty u^0 u u_x^0 + v^0 u u_y^0 + \nu/2\, u_y^2 \mathrm{d}y\, \mathrm{d}x. \tag{8}$$

(d) *Venkateshwarlu* [16]

$$I_v = \int_0^l \int_0^\infty 1/2\, (uuu_x^0 - uu^0 u_x) + 1/2\, (vuu_y^0 - vu^0 u_y)$$

$$+ \nu/2\, u_y^2\, \mathrm{d}y\, \mathrm{d}x + 1/2 \int_0^\infty uu^0 u \Big|_{x=l} \mathrm{d}y. \tag{9}$$

We shall now establish that for a given choice of trial function the various local potential formulations (6) to (9) will lead to identical end results. Performing the variational process by the Euler–Lagrange equations in the functionals (6) to (9), and substituting the auxiliary conditions

$$u = u^0, \quad v = v^0, \tag{10}$$

and the boundary conditions we get

$$\delta I = \int_0^l \int_0^\infty e(x, y)\, \mathrm{d}u\, \mathrm{d}y\, \mathrm{d}x, \tag{11}$$

where

$$e\ (x, y) = \nu \frac{\partial^2 u}{\partial y^2} - u \frac{\partial u}{\partial x} - v \frac{\partial u}{\partial y}. \tag{12}$$

The form of the velocity profile, generally, employed is

$$u = UF(\eta), \quad \eta = y/\delta\,(x), \tag{13}$$

The procedure is to substitute (13) in the desired local potential formulation and then minimising the functional with respect to δ. This minimisation leads to a differential equation whose solution gives the required results. This process is equivalent to rewriting (11) as

$$\delta I = \int_0^l \int_0^\infty e \varphi\, \mathrm{d}\delta\, \mathrm{d}y\, \mathrm{d}x, \quad \varphi = \partial u/\partial \delta. \tag{14a, 14b}$$

The problem resulting from (14) is

$$\int_0^\infty \varphi e\, \mathrm{d}y = 0. \tag{15}$$

The relation (15) is the moment of boundary layer equation (2) by a function $\phi = \partial u/\partial \delta$ whereas the Karman's integral corresponds to $\phi = 1$. Each of the local potential formulations (6) to (9) is equivalent to solving (15) and thus the end results from various formulations are the same.

As long as one has to start with an assumed form of velocity profile with some free variables, one might determine the free variables by minimising the error in the least square sense. For example if u and v are the approximating functions satisfying the boundary conditions, involves certain unknown free parameters, then the residual or error in the boundary layer equation (2) is

$$e\,(x, y, \nu) = \nu u_{yy} - u u_x - v u_y + U U_x. \tag{16a}$$

The total error is defined as the integral of e^2 over the whole range of the relevant variables

$$E = \int_0^l \int_0^\infty e^2\, \mathrm{d}y\, \mathrm{d}x. \tag{16b}$$

The functional E may be minimized with respect to free parameters by employing the Euler–Lagrange equations which results in certain equations whose solution leads to the full determination of the velocity profile. The procedure is illustrated by solving several boundary layer problems described in the next section.

3. Boundary layer on a flat plate

The boundary layer on a flat plate has been extensively studied [10, 14]. Using the local potential method the problem has been studied by [12, 16] and the variational principle by [17]. We study here the problem by the least square method. The approximating function, which satisfies the boundary conditions, is generally taken [10, 12, 14, 16] of the form

$$u = Uf'(\eta), \quad \eta = y/\delta(x). \tag{17}$$

Substituting (17) in the boundary layer equation (2), the residual (16a) is given by

$$e = \frac{\nu U}{\delta^2}\left(f''' + \frac{U\delta\delta'}{\nu} ff''\right). \tag{18}$$

The total error E, defined by (16b), is given by

$$E = \nu^2 U^2 \int_0^l \left[\frac{J_1}{\delta^3} + \frac{2U\delta'}{\nu\delta^2} J_2 + \frac{U^2\delta'^2}{\nu^2\delta} J_3\right] dx, \tag{19}$$

where

$$J_1 = \int_0^\infty f'''^2 \, d\eta, \quad J_2 = \int_0^\infty ff''f''' \, d\eta,$$

and

$$J_3 = \int_0^\infty f^2 f''^2 \, d\eta, \tag{20}$$

are certain constants depending upon the choice of profile. Minimising the total error E with respect to δ by the Euler–Lagrange equation

$$\frac{\partial E}{\partial \delta} - \frac{d}{x}\frac{\partial E}{\partial \delta'} = 0, \tag{21}$$

we get the second order nonlinear ordinary differential equation

$$2\delta^3\delta'' - \delta^2\delta'^2 + \frac{3\nu^2}{U^2}\frac{J_1}{J_3} = 0. \tag{22}$$

Assuming the following solution

$$\delta(x) = A x^n, \tag{23}$$

the values of n and A obtained by substituting (23) in (22) are

$$n = 1/2, \tag{24a}$$

$$A = (\nu/U)^{1/2} (4J_1/J_3)^{1/4}. \tag{24b}$$

To test the present method the five different choices of velocity profile

$$(a)\ f'(\eta) = 3/2\eta - 1/2\eta^3\ ; \quad (b)\ f'(\eta) = \sin \pi/2\ \eta\ ;$$
$$(c)\ f'(\eta) = 2\eta - 2\eta^3 + \eta^4\ ; \quad (d)\ f'(\eta) = 1 - e^{-\eta};$$
$$(e)\ f'(\eta) = Erf(\eta) \tag{25}$$

usually employed in the literature are considered. For these profiles the values of J_1, J_2, J_3 and A are given in table 1. The predictions for the skin friction coefficient C_f, displacement thickness δ^* and momentum thickness δ^{**} are given in table 2. The available results based on the Karman–Pohlhausen method [3], the local potential method [12, 16] and Vujanovic and Djukic method [16, 17] are also given in table 2 for comparison. Table 2 shows that the present predictions based on the least square method are invariably better.

4. Hiemenz flow

The Hiemenz flow near the two-dimensional stagnation point ($U = ax$) has been studied by Doty and Blick [1] through local potential formulation of Glansdorff and Prigogine [2]. The velocity profile employed by [1] is

$$u = Uf'(\eta), \quad f' = 1 - e^{-\eta}, \tag{26a, b}$$
$$\eta = y/\delta, \qquad \delta = (2\nu/a)^{1/2}. \tag{26c, d}$$

For this case the total error E, defined by (16), obtained from boundary layer equation (2) is

$$E = \int_0^l \int_0^\infty \frac{\nu U^2}{\delta^3}\left[f''' + \frac{\delta^2 a}{\nu}\,(ff'' - f'^2 + 1)\right]^2 d\eta\, dx. \tag{27}$$

Table 1. Flat plate boundary layer : Values of constants J_1, J_2, J_3 and A for various profiles for minimum error solution.

Profile $f'(\eta)$	J_1	J_2	J_3	$A(U/\nu)^{1/2}$
1. $3/2\eta - 1/2\eta^3$	3	$\frac{33}{128}$	$\frac{629}{22880}$	4·57084
2. $\sin \pi/2\eta$	$\frac{\pi^4}{32}$	$-\frac{\pi}{12}$	$\frac{21\pi - 64}{24\pi}$	4·64421
3. $2\eta - 2\eta^3 + \eta^4$	$\frac{24}{5}$	$-\frac{50629}{11550}$	$\frac{386636}{19144125}$	5·55276
4. $1 - e^{-\eta}$	$\frac{1}{2}$	$-\frac{1}{12}$	$\frac{1}{18}$	2·44948
5. Erf (η)	0·797884	− 0·156316	0·0341957	3·108186

Table 2. Flat plate boundary layer: Comparison of skin friction C_f, displacement thickness δ^* and momentum thickness δ^{**} predicted by various methods.

Profile	$1/2\, C_f \sqrt{R_x}$				$\delta^* (U/\nu x)^{1/2}$				$\delta^{**} (U/\nu x)^{1/2}$			
	Present	Integral	Local potential	D.V. method	Present	Integral	Local potential	D.V. method	Present	Integral	Local potential	D.V. method
1. $\frac{3}{2}\eta - \frac{1}{2}\eta^3$	0·328	0·323	...	...	1·714	1·74	...	...	0·637	0·646	...	...
2. $\sin \frac{\pi\eta}{2}$	0·338	0·328	...	...	1·688	1·743	...	...	0·634	0·654	...	...
3. $2\eta - 2\eta^3 + \eta^4$	0·360	0·343	...	...	1·66	1·75	...	...	0·652	0·685	...	...
4. $1 - e^{-\eta}$	0·408	0·5	0·471	0·304	2·449	2·0	2·121	3·281	1·225	1·0	0·943	0·608
5. $Erf\ (\eta)$	0·363	0·363	0·352	0·166	1·75	1·75	1·865	3·838	0·726	0·726	0·704	0·332
Exact result	0·332				1·720				0·664			

Minimising the total error E with respect to δ we get

$$\delta = (\nu/a)^{1/2} \quad (J_1/J_3)^{1/4}, \tag{28}$$

where

$$J_1 = \int_0^\infty f'''^2 \, d\eta, \quad J_2 = \int_0^\infty f'' \, (ff'' - f'^2 + 1) \, d\eta, \tag{29a, b}$$

$$J_3 = \int_0^\infty (ff'' - f'^2 + 1)^2 \, d\eta. \tag{29c}$$

For the profile (26b) the values of integral are

$$J_1 = 1/2, \quad J_2 = 3/4, \quad J_3 = 5/4, \tag{30}$$

and the boundary layer thickness δ is given by

$$\delta = (2/5)^{1/4} \, (\nu/a)^{1/2}. \tag{31}$$

The coefficient of skin friction C_f is given by

$$1/2 \; C_f \sqrt{R_x} = 1 \cdot 2574. \tag{32}$$

The value predicted by Doty and Blick [1] from the local potential method is 1·4142 and the exact result [14] is 1·2326. The prediction based on the local potential method overestimates the skin friction by 14% and the present result by 2%. Thus the predictions based upon the least square method are better than that given by local potential.

5. Boundary layer on a moving sheet

The boundary layer on a continuous sheet moving with a constant velocity U_0 in an stationary ambient fluid was studied by Sakiadis [11]. Let the approximating function be of the form

$$u = U_0 \, f'(\eta), \quad \eta = y/\delta\,(x), \tag{33}$$

which satisfies the boundary conditions

$$f(o) = 0, \quad f'(o) = 1, \quad f'(\infty) = 0, \tag{34}$$

for a moving sheet. Using the profile (33) the total error (16) is given by

$$E = \nu^2 \, U_0^2 \int_0^1 \frac{J_1}{\delta^3} + \frac{2U_0 \, \delta'}{\nu \delta^2} \, J_2 + \frac{U_0^2 \, \delta'^2}{\nu^2 \, \delta} \, J_3 \, dx, \tag{35}$$

where the forms of J_1, J_2 and J_3 are given by (21). Minimizing E with respect to δ by employing the Euler–Lagrange equations we get

$$\delta = (4J_1/J_2)^{1/4} \, (\nu x/U_0)^{1/2}. \tag{36}$$

If we choose a profile

$$f'(\eta) = Erfc(\eta), \tag{37}$$

which satisfies the boundary conditions (34) then the constants (21) are given by

$$J_1 = 0{\cdot}79788, \quad J_2 = -0{\cdot}24263 \quad \text{and } J_3 = 0{\cdot}07735, \tag{38}$$

and the boundary layer thickness δ is given by

$$\delta = 2{\cdot}5344\,(\nu x/U_0)^{1/2}. \tag{39}$$

The coefficient of skin friction using (39) may be written as

$$1/2\ C_f\sqrt{R_x} = 0{\cdot}4452 \tag{40}$$

The Karman–Pohlhausen method for the profile (37) gives a value of 0·4318 and the exact result [11] is 0·4437. Thus here again the predictions based upon the least square method are more accurate when compared with Karman–Pohlhausen method.

6. Boundary layer in non-Newtonian fluids

The boundary layer on a flat plate in non-Newtonian fluids has been studied by several workers and a good review has been given by Skelland [15]. We consider here the power law fluids where shear stress τ is given by

$$\tau = K(\partial u/\partial y)^n, \tag{41}$$

where K and n are certain constants of the fluid.

For the boundary layer on a flat plate the approximating profile be of the form

$$u = Uf'(\eta), \quad \eta = y/\delta(x). \tag{42}$$

The residual in the boundary layer equation (2) is

$$e = \frac{KU^2}{\rho\,\delta^{n+1}}\left[(f''^n)' + \rho\,U^{2-n}\,\frac{\delta\,\delta'}{K}\,ff''\right]. \tag{43}$$

The total error E, defined by (16), is given by

$$E = \frac{K^2U^{2n}}{\rho^2}\int_0^l \frac{J_1}{\delta^{2n+1}} + \frac{2\rho U^{2-n}}{K}\,\frac{\delta'}{\delta^{2n+1}}\,{}_2 + \frac{\rho^2\,U^{4-2n}\,\delta'^2}{K^2\delta^{2n}}\,J_3\,dx. \tag{44}$$

where

$$J_1 = \int_0^\infty [(f''^n)']^2\,d\eta, \quad J_2 = \int_0^\infty ff''(f''^n)'\,d\eta \text{ and } J_3 = \int_0^\infty f^2f''^2\,d\eta. \tag{45}$$

Minimising the total error E with respect to δ by employing the Euler–Lagrange equation we get a second order ordinary differential equation

$$2\delta^{2n+1}\delta'' - \delta^{2n}\delta'^2 + (2n+1)\,\frac{K^2U^{2n-4}}{\rho^2}\,\frac{J_3}{J_1} = 0. \tag{46}$$

Assuming the solution of the form

$$\delta = CA\, x^p, \tag{47}$$

equation (46) leads to

$$p = \frac{1}{1+n}, \quad C = \left(\frac{K}{\rho U^{2-n}}\right)^{1/1+n}, \tag{48a, b}$$

$$A = \left[(n+1)^2 \frac{J_1}{J_3}\right]^{1/2n+4}. \tag{48c}$$

Now assuming the cubic velocity profile [15]

$$f'(\eta) = 1/2\,(3\eta - \eta^3), \tag{49}$$

which satisfies the boundary conditions, the integrals (45) are given by

$$\left.\begin{aligned} J_1 &= (9/4)^n\, n^2 \sqrt{\pi}\, \Gamma(2n-1)/\, \Gamma\tfrac{1}{2}(4n+1), \\ J_2 &= -\,3/8\,(3/2)^n \frac{n(3n+8)}{(n+1)(n+2)(n+3)}, \\ J_3 &= \frac{629}{22880}, \end{aligned}\right\} \tag{50}$$

and A is given by

$$A = \left[\frac{22880}{629}\left(\frac{9}{4}\right)^n \sqrt{\pi}\, n^2 (n+1)^2 \,.\, \frac{\Gamma(2n-1)}{\Gamma\left(\frac{4n+1}{2}\right)}\right]^{1/2n+4}. \tag{51}$$

The skin friction coefficient C_f is given by

$$1/2\; C_f\,(R_{ex})^{1/1+n} = (3/2A)^n, \tag{52}$$

where R_{ex} is a Reynolds number defined by

$$R_{ex} = \frac{\rho U^{2-n}}{K} x^n. \tag{53}$$

The value C_f obtained from the Karman method [15] is

$$1/2\; C_f\,(R_{ex})^{1/1+n} = \left[\frac{39}{280}\left(\frac{1\cdot 5}{n+1}\right)\right]^{n/n+1}. \tag{54}$$

A comparison of the present result (52) with the exact results [15] and the result (54) from Karman–Pohlhausen method is given in table 3 for various n values. The table shows that the present results are better than (54) when compared with exact results.

7. Conclusions

(a) All the local potential methods [2, 4, 6, 16] for a given choice of trial function lead to identical end results, as all the formulations are equivalent to the moment of boundary layer equation by a function $\phi = \partial u/\partial\delta$.

(b) A method has been developed to study the minimum error solution of boundary layer equations in the least square sense by employing the Euler–Lagrange

Table 3. Boundary layer in non-Newtonian fluids : Comparison of skin friction with available results.

n	A	$1/2\, C_f\, (R_{ex})^{1/1+n}$		
		Present	Integral method [15]	Exact [15]
0·1	1·8214	0·9807	0·860	0·969
0·2	3·1000	0·8649	0·7474	0·8725
0·3	4·4657	0·7209	0·6555	0·7325
0·5	3·9447	0·6166	0·518	0·5755
1·0	4·5708	0·3282	0·323	0·3321
1·5	3·9159	0·2371	0·2255	0·2189
2·0	3·5555	0·1780	0·169	0·1612
2·5	3·3039	0·1389	0·133	0·1226
3·0	3·1129	0·1119	0·109	0·0971
4·0	2·8381	0·0780	0·079	0·0678
5·0	2·6484	0·0583	0·061	0·0511

equations. A class of problems in the boundary layer theory has been studied for a variety of trial functions and it is found that the present method gives invariably the best results when compared with the results of other approximate methods like Karman–Pohlhausen, local potential and Vujanovic and Djukic.

Acknowledgement

I am thankful to Prof. M S Deshpande for some helpful discussions.

References

[1] Doty R T and Blick E F 1973 *AIAA J.* **11** 880
[2] Glansdorff P and Prigogine I 1964 *Physica* **30** 331
[3] Goldstein S 1965 Modern developments in fluid dynamics (New York : Dover) Vol. **1**
[4] Hiroaka M and Tanaka K 1968 *Mem. Fac. Engg. Technol. Kyoto Univ.* Part 4 **30** 397
[5] Hsu C C 1975 *J. Fluid Mech.* **69** 783
[6] Lambermont J and Lebon G 1972 *Ann. Phys.* **7** 15
[7] Lebon G and Lambermont J 1973 *J. Chem. Phys.* **59** 2929
[8] MacDonald D A 1974 *Int. J. Heat Mass Transfer* **17** 393
[9] Prigogine I and Glansdorff P 1965 *Physica* **31** 1243
[10] Rosenhead L 1963 *Laminar boundary layers* (Oxford : University Press)
[11] Sakiadis B C 1961 *A.I.Ch.E. J.* **7** 26
[12] Schechter R S 1962 *The variational methods in engineering* (New York : McGraw-Hill)
[13] Schechter R S 1966 Non equilibrium thermodynamics : *Variational techniques and stability* (eds) R J Donnelly, R Hermann and I Prigogine (Chicago : University Press)
[14] Schlichting H 1968 *Boundary layer theory* (New York : McGraw-Hill)
[15] Skelland A H 1967 *Non-Newtonian flow and heat transfer* (New York : John Wiley)
[16] Venkateshwarlu P 1978 *A critical study of variational formulations in fluid flow and heat transfer and some applications*, Ph.D. thesis, Indian Institute of Science, Bangalore
[17] Vujanovic B and Djukic D J 1971 *Publs. Inst. Maths.* **11** 73

Proc. Indian Acad. Sci. (Math. Sci.), Vol. 91, Number 3, November 1982, pp. 195–199.

Complementary variational principles for poiseuille flow of an Oldroyd fluid

M A GOPALAN
Department of Mathematics, National College, Trichy 620 001, India

MS received 7 October 1981 ; revised 26 June 1982

Abstract. The complementary variational principles are given for the poiseuille flow of an Oldroyd fluid by taking the pressure gradient to be exponentially increasing with time and the bounds on the flux are obtained.

Keywords. Variational principles ; complementary principles ; bounds on flux ; poiseuille flow; Oldroyd fluid.

1. Introduction

Complementary variational principles for steady and unsteady poiseuille flow of a Newtonian fluid have been given earlier [2]. Non-Newtonian fluids are also important in the present day technology. Soundalgekar [4–6] has discussed the flow of non-Newtonian fluids through pipes by taking the pressure gradient to be either exponentially increasing or decreasing. Bhatnagar (1) has considered the fluctuating flow of an Oldroyd fluid through a straight circular tube. The aim of the present paper is to develop complementary variational principles for poiseuille flow of an Oldroyd fluid. It is seen that these extremum principles provide upper and lower bounds on the amplitude of the mass flow, when the axial pressure gradient increases exponentially with time.

2. Mathematical formulation

The equations governing the elastico-viscous fluid model as proposed by Oldroyd [3] consist of stress-strain rate law

$$p_{ik} = -pg_{ik} + p'_{ik}, \tag{1}$$

$$\left(1 + \lambda_1 \frac{D}{Dt}\right) p'_{ik} = 2\mu\left(1 + \lambda_2 \frac{D}{Dt}\right) d_{ik}, \tag{2}$$

where p_{ik} is the stress tensor, p an arbitrary isotropic pressure, g_{ik} is the metric tensor and d_{ik} is the rate of straintensor. μ, λ_1, λ_2 are material constants.

Consider fully developed poiseuille flow of an Oldroyd fluid in a long straight pipe of arbitrary cross-section S with X-axis coinciding with the axis of the pipe. An unsteady rectilinear flow through the pipe can be represented by the velocity field $(0, 0, u^*(y, z, t))$. The momentum equation determining $u^*(y, z, t)$ has the form

$$\rho\left(1+\lambda_1 \frac{\partial}{\partial t}\right)\frac{\partial u^*}{\partial t} = -\left(1+\lambda_1\frac{\partial}{\partial t}\right)\frac{\partial p}{\partial x} + \mu\left(1+\lambda_2\frac{\partial}{\partial t}\right)\nabla^2 u^*. \tag{3}$$

If $\dfrac{\partial p}{\partial x} = -\gamma \exp(n^2 t)$, $u^* = u \exp(n^2 t)$.

Equation (3) reduces to

$$\nabla^2 u = Mu - (\gamma N/\mu), \tag{4}$$

subject to the condition that

$$u = 0 \text{ on boundary } C, \tag{5}$$

where $M = (n^2 N/\nu)$, $N = (1 + \lambda_1 n^2/1 + \lambda_2 n^2)$.

3. Complementary principles

Equation (4) can be expressed in the form

$$\text{grad } U = \Phi,\ U = 0 \text{ on } C, \tag{6}$$

$$-\text{div } \Phi = (\gamma N/\mu) - MU, \tag{7}$$

where Φ is a vector having components in the y and z directions. Consider the functional

$$I(U, \Phi) = \int_s \left(\tfrac{1}{2}\Phi\cdot\Phi + (\gamma/\mu) NU - (M/2) U^2 - \Phi\cdot\nabla U\right) dS$$

$$+ \int_C U\Phi\cdot\hat{n}\, dC. \tag{8}$$

$$= \int_S \left(\tfrac{1}{2}\, \Phi\cdot\Phi + (\gamma/\mu)\, NU - (M/2)\, U^2 + U \text{ div } \Phi\right) dS. \tag{9}$$

The extremals of this functional with $U = 0$ on C lead to (6) and (7). The exact solution is denoted by $U = u$, $\Phi = \phi$. The complementary principles are constructed from (6) to (9) as follows :

First choose a trial function U satisfying (6). Then (8) gives

$$G(U) = \int_S \left[(\gamma/\mu)\, NU - (M/2)\, U^2 - (1/2)\nabla U\cdot\nabla U\right] dS. \tag{10}$$

Next choose another trial function Φ satisfying (7). Then (9) gives

$$J(\Phi) = \int_S \left[(1/2)\, \Phi\cdot\Phi + (1/2)\, M^{-1}\, (\gamma N/\mu + \text{div } \Phi)^2\right] dS. \tag{11}$$

The functionals $G(U)$ and $J(\Phi)$ provide lower and upper bounds to the functional $I(u,\phi)$ that is,

$$G(U) \leqslant G(u) = I(u,\phi) = J(\phi) \leqslant J(\Phi) \tag{12}$$

From (4), the flux Q is given by

$$Q = (2\mu/\gamma N)\ I\ (u, \phi) \tag{13}$$

Thus we obtain

$$(2\mu/\gamma N)\ G\,(U) \leqslant Q \leqslant (2\mu/\gamma N)\ J\,(\Phi) \tag{14}$$

4. Calculation of upper and lower bounds

Let $f(y, z) = 0$ be the equation of the bounding curve C. Choose U as

$$U = (\gamma N/\mu)\ \alpha\ f\ (y, z), \tag{15}$$

where α is a parameter and the boundary condition is satisfied. Substituting in (10) we get

$$G\,(U) = \left(\frac{\gamma N}{\mu}\right)^2 \left[\alpha \int_S f \mathrm{d}S - \frac{1}{2}\alpha^2 \int_S M f^2\, \mathrm{d}S - \frac{1}{2}\alpha^2 \int_S \nabla f \cdot \nabla f \mathrm{d}S\right]. \tag{16}$$

The extremals of (16) are obtained by setting $\partial G/\partial\alpha = 0$. This gives

$$\alpha = \int_S f \mathrm{d}S \Big/ \triangle,\ \frac{\partial^2 G}{\partial \alpha^2} < 0, \tag{17}$$

where

$$\triangle = \int_S (\nabla f \cdot \nabla f + M f^2)\, \mathrm{d}S.$$

Therefore,

$$G\,(U) = \frac{1}{2}\left(\frac{\gamma N}{\mu}\right)^2 \left(\int_S f \mathrm{d}S\right)^2 \triangle^{-1}, \tag{18}$$

is the required lower bound and the greatest lower bound is the exact solution. Again, choose

$$\Phi = -\frac{\gamma N}{\mu}\,[\beta y \hat{j} + (1-\beta)\, z \hat{k}] + \frac{\gamma}{\mu}\, M\delta\,[h\,(y, z)\,\hat{j} + g\,(y, z)\,\hat{k}]$$

where

$$\frac{\gamma\delta}{\mu}\left(\frac{\partial h}{\partial y} + \frac{\partial g}{\partial z}\right) = U. \tag{18a}$$

so that (7) is satisfied. Substituting in (11) we get

$$J\,(\Phi) = \frac{1}{2}\,\frac{\gamma^2}{\mu^2}\,\{M\delta^2\,[P_5 + M\,(P_1 + P_3)] + \beta^2 N^2 I_3 + N^2\,(1-\beta)^2 I_2 - 2MN\delta\,[\beta P_2 + (1-\beta) P_4]\} \tag{19}$$

where

$$P_1 = \int_S h^2\, dS,\ P_3 = \int_S g^2\, dS,\ I_3 = \int_S y^2\, dS,\ I_2 = \int_S z^2\, dS,$$

$$P_2 = \int_S yh\, dS,\ P_4 = \int_S zg\, dS,\ P_5 = \int_S \left(\frac{\partial h}{\partial y} + \frac{\partial g}{\partial z}\right)^2 dS.$$

The extremals of (19) are given by $\partial J/\partial\beta = 0$, $\partial J/\partial\delta = 0$.

This gives

$$\left.\begin{aligned} &\delta = N(P_2 I_2 + P_4 I_3)/\triangle_1 \\ &\beta = I_2 P_5 + M\,[I_2(P_1 + P_3) + P_4(P_2 - P_4)]/\triangle_1 \\ &J_{\delta\delta}\, J_{\beta\beta} - J_{\beta\delta}^2 > 0,\quad J_{\beta\beta} > 0,\quad J_{\delta\delta} > 0, \end{aligned}\right\} \tag{20}$$

where

$$\triangle_1 = (I_2 + I_3)\ [M(P_1 + P_3) + P_5] - M(P_2 - P_4)^2,\ \ J_{\beta\beta} = \frac{\partial^2 J}{\partial\beta^2}$$

From (19) and (20) we get $J\ (\Phi)$ as the required upper bound and the least upper bound is the exact solution.

5. Application

Circle : $y^2 + z^2 = a^2$

In parametric co-ordinates $y = ar\cos\theta$, $z = ar\sin\theta$, $0 \leqslant r \leqslant 1$, $0 \leqslant \theta \leqslant 2\pi$, the boundary C corresponds to $r = 1$. Take $f(y, z) = 1 - r^2$.
From (17) and (18)

$$a = (3a^2/(12 + 2Ma^2)),$$

$$G(U) = \frac{1}{2}\left(\frac{\gamma N}{\mu}\right)^2 [3\pi a^4 \big/ (24 + 4Ma^2)]. \tag{21}$$

Again take $U = (\gamma\delta/\mu)\,(1 - r^2)$ and from (18a)

$$h = y(2 - r^2)/4,\quad g = z(2 - r^2)/4$$

From (19) and (20)

$$\beta = 1/2,\quad \delta = 16Na^2/(64 + 11Ma^2),$$

$$J(\Phi) = \frac{\pi a^4}{2}\left(\frac{\gamma N}{\mu}\right)^2\left[\frac{1}{8} - \frac{4Ma^2}{3\,(64 + 11\,Ma^2)}\right] \tag{22}$$

Therefore from (14), (21) and (22) we have

$$\frac{\gamma N}{\mu}\left(\frac{3\pi a^4}{24 + 4Ma^2}\right) \leqslant Q \leqslant \frac{\gamma N}{\mu}\pi a^4\left(\frac{1}{8} - \frac{4Ma^2}{3\,(64 + 11Ma^2)}\right) \tag{23}$$

In conclusion, the bounds on the flux can be obtained for other simple geometrics. Further, the bounds for Maxwell fluid [4] second order fluid [5] and Walters fluid [6] are obtained by appropriately choosing the two fluid parameters λ_1 and λ_2. If both the parameters are taken as zero, the results for classical viscous fluids are obtained.

Using the binomial theorem and neglecting the terms of order M^2 in (23), it is seen that the upper and lower bounds coincide. In particular, neglecting the terms of order M, the amplitude of the mass flow for a second order fluid is found to be in agreement with equation (15) of [5].

References

[1] Bhatnagar R K 1975 *Appl. Sci. Res.* **30** 241

[2] Jamal N G, Kamala V, Prabhakara Rao G V and Nigam S D 1979 *J. Math. Phys. Soc.* **13** 469

[3] Oldroyd J G 1950 *Proc. R. Soc. (London)* **A200** 523

[4] Soundalgekar V M 1965 *AIAA J.* **3** 968

[5] Soundalgekar V M 1972 *Indian J. Phys.* **46** 250

[6] Soundalgekar V M 1973 *Chem. Engg. Sci.* **28** 654

Proc. Indian Acad. Sci. (Math. Sci.), Vol. 91, Number 3, November 1982, pp. 201–210.

Torsional wave propagation in a finite inhomogeneous cylindrical shell under time-dependent shearing stress

K VENKATESWARA SARMA
Department of Mathematics, PSG College of Arts and Science, Coimbatore 641 014, India

MS received 3 August 1981 ; revised 1 July 1982

Abstract. Torsional wave motion in a finite hollow cylinder of piezoelectric material of (622) crystal class, under a time-dependent mechanical boundary condition is investigated. The inhomogeneity is restricted to the variations of density and other physical constants of the medium as a certain power of the radial distance. The expressions for the displacement and the electric potential of the present solution are compared with those under time-dependent electric boundary condition. Numerical values of the roots of the frequency equation for β-quartz are presented.

Keywords. Piezoelectricity ; torsional wave ; shearing stress ; inhomogeneous cylindrical shell.

1. Introduction

Torsional oscillations of a finite inhomogeneous piezoelectric cylindrical shell under a time-dependent electric potential on the boundary have been discussed earlier [3]. The present paper investigates the torsional oscillations of a similar shell under the influence of a time-dependent shearing stress on the outer cylindrical surface of the shell. For numerical study, the roots of the frequency equation for β-quartz are computed on the digital computer IBM 370/155. The first few roots are listed in table 1.

2. Basic equations and the boundary conditions of the problem

The axis of the shell is along the z-axis of the cylindrical coordinate system (r,θ,z) and the ends of the shell lie on the planes $z = 0$ and $z = L$, L being the length of the shell. The curved surfaces have the radii a and b $(a < b)$. The torsional wave motion is characterized by a displacement $v(r,z,t)$ solely in the θ direction. ϕ (r,z,t) denotes the electric potential. Both v and ϕ are independent of θ due to axi-symmetry of the problem. The constitutive relations appropriate for a piezoelectric material of (622) crystal class are, [3],

$$T_{\theta z} = c_{44} v,_z + e_{14} \phi,_r \; ; \quad D_r = e_{14} v,_z - \epsilon_{11} \phi,_r$$

$$T_{r\theta} = c_{66} \left(v,_r - \frac{v}{r} \right) ; \quad D_z = - \epsilon_{33} \phi,_z \tag{1}$$

where T's are the components of stress, D's are the components of electric displacement vector, and c, e and ϵ's denote respectively the elastic, piezoelectric and dielectric constants. A comma as a subscript in the above relations denotes partial differentiation with respect to the indicated variable. The inhomogeneity of the material is assumed to be such that the physical constants vary as the $2N$th power of the radial distance r. Thus

$$[\rho, c_{44}, c_{66}, e_{14}, \epsilon_{11}, \epsilon_{33}] = [s, C_{44}, C_{66}, E_{14}, E_{11}, E_{33}] \, r^{2N}, \tag{2}$$

where ρ is the density of the material.

In the absence of body forces the equations of motion and the electrostatic charge are respectively [3],

$$T_{r\theta},_r + T_{z\theta},_z + \frac{2}{r} T_{r\theta} = \rho v,_{tt}$$

$$D_{r},_r + D_{z},_z + \frac{1}{r} D_r = 0. \tag{3}$$

Substituting relations (1) and (2) in (3), the equations for v and ϕ are

$$C_{44} v,_{zz} + C_{66} \nabla_N^2 (v) + E_{14} \phi,_{rz} = s v,_{tt}$$

$$E_{14} \left[v,_{rz} + \frac{2N+1}{r} v,_z \right] - E_{11} \left[\phi,_{rr} + \frac{2N+1}{r} \phi,_r \right] - E_{33} \phi,_{zz} = 0 \tag{4}$$

where

$$\nabla_N^2 (\;) = \left[\frac{\partial^2}{\partial r^2} + \frac{2N+1}{r} \frac{\partial}{\partial r} - \frac{2N+1}{r^2} \right] (\;).$$

All boundary surfaces are free from any applied potential ; the outer cylindrical surface of the shell is subjected to a time-dependent shearing stress, and the end faces of the shell are so fixed that no point on them suffers any displacement. Thus

$$T_{z\theta} = 0 \text{ for } z = O, L,$$

$$\phi = 0 \text{ for } z = O, L \text{ and for } r = a, b, \tag{5}$$

and

$$T_{r\theta} = C_{66} \, p(z, t) \text{ for } r = b,$$

$$= 0 \qquad \text{for } r = a.$$

3. Solution of the boundary value problem

To solve the boundary value problem (4) and (5), the expressions for v, ϕ and $p(z,t)$ are assumed in the form

$$\phi = \sum_{m=1}^{\infty} \{\Phi_m(r, t) + F_m(r, t)\} \sin \alpha_m z,$$

$$v = \sum_{m=1}^{\infty} \{V_m(r, t) + G_m(r, t)\} \cos \alpha_m z \tag{6}$$

and

$$p = \sum_{m=1}^{\infty} P_m(t) \cos(\alpha_m z)$$

where

$$\alpha_m = m\pi/L.$$

Substitution of the above expressions in (4) and (5) result in the following equations.

$$-\alpha_m^2 C_{44} V_m + C_{66} \nabla_N^2 (V_m) + \alpha_m E_{14} \Phi_{m,r} - s V_{m,tt} \tag{7}$$
$$= sG_{m,tt} + \alpha_m^2 C_{44} G_m - C_{66} \nabla_N^2 G_m - \alpha_m E_{14} F_{m,r},$$

$$E_{14}\alpha_m\left(V_{m,r} + \frac{2N+1}{r} V_m\right) + E_{11}\left(\Phi_{m,rr} + \frac{2N+1}{r}\Phi_{m,r}\right) - E_{33}\alpha_m^2 \Phi_m$$
$$= -E_{14}\alpha_m\left(G_{m,r} + \frac{2N+1}{r} G_m\right)$$
$$-E_{11}\left(F_{m,rr} + \frac{2N+1}{r} F_{m,r}\right) + E_{33}\alpha_m^2 F_m. \tag{8}$$

$T_{z\theta} = 0$ is satisfied automatically from (1) and (6). Other conditions give

$$\Phi_m(r, t) = -F_m(r, t) \text{ for } r = a, b,$$
$$V_{m,r} - \frac{1}{r} V_m = \frac{1}{r} G_m - G_{m,r} \text{ for } r = a \tag{9}$$
$$= P_m(t) + \frac{1}{r} G_m - G_{m,r} \text{ for } r = b.$$

To remove the time-dependency in the boundary conditions (9), the Fourier coefficients G_m and F_m are so prescribed as to satisfy

(i) $\nabla_N^2 G_m - \lambda_m^2 G_m = 0$ where $\lambda_m^2 = \alpha_m^2 \dfrac{C_{44}}{C_{66}}$,

(ii) The right side of equation (8) vanishes,

(iii) $\dfrac{1}{r} G_m - G_{m,r} = 0$ for $r = a$, (10)

(iv) $P_m(t) + \dfrac{1}{r} G_m - G_{m,r} = 0$ for $r = b$,

and

(v) $F_m(r,t) = 0$ for $r = a, b$.

Seeking product type solutions for G_m and F_m in the forms

$$G_m(r, t) = g_m(r)\, P_m(t), \qquad F_m(r,t) = f_m(r)\, P_m(t). \tag{11}$$

the set of equations (10) yields

$$g_m^{11} + \frac{2N+1}{r} g_m^1 - \left(\lambda_m^2 + \frac{2N+1}{r}\right) g_m = 0 \tag{12}$$

$$g_m^1(r) - \frac{1}{r} g_m(r) = 0 \; for \; r = a$$
$$= 1 \text{ for } r = b. \tag{13}$$

The solution of the system (12) and (13) is

$$g_m(r) = r^{-N}\left[A_m I_{N+1}(\lambda_m r) + B_m K_{N+1}(\lambda_m r)\right], \tag{14}$$

where $I_{N+1}(\;)$ and $K_{N+1}(\;)$ are modified Bessel functions. The constants A_m and B_m are given by

$$A_m = -\frac{y_1}{x_1 y_2 - x_2 y_1}\,; \quad B_m = \frac{x_1}{x_1 y_2 - x_2 y_1},$$

where

$$x_1 = I_{N+2}(\lambda_m a)\,; \quad y_1 = K_{N+2}(\lambda_m a),$$
$$x_2 = \lambda_m b^{-N} I_{N+2}(\lambda_m b)\,; \quad y_2 = \lambda_m b^{-N} K_{N+2}(\lambda_m b). \tag{15}$$

In view of the result in (11), the result 10 (ii) simplifies to

$$f_m^{11} + \frac{2N+1}{r} f_m^1 - \beta_m^2 f_m = h_m(r), \tag{16}$$

where

$$\beta_m^2 = \alpha_m^2 \frac{E_{33}}{E_{11}} \text{ and } h_m(r) = \frac{-\alpha_m E_{14}}{E_{11}}\left(g_m^1 + \frac{2N+1}{r} g_m\right),$$

$h_m(r)$ is known through (14) and (15). Further, from 10 (V),

$$f_m(a) = 0 = f_m(b). \tag{17}$$

To solve the boundary value problem given by (16) and (17), the method of Green's functions [1], is used. The condition for the existence of Green's function is satisfied ; vide § 4·5, equation (4·7) in, [4] ; since

$I_N(\beta_m a)\, K_N(\beta_m b) \neq I_N(\beta_m b)\, K_N(\beta_m a)$. Equation (16) is rewritten as

$$(r^{2N+1} f_m^1)^1 - \beta_m^2 r^{2N+1} f_m = r^{2N+1} h_m. \tag{18}$$

Let $u_m(r)$ and $v_m(r)$ be the solutions of the homogeneous equation corresponding to (18) and satisfying the conditions

$$u_m(a) = 0\,; \quad u_m^1(a) = 1,$$
$$v_m(b) = 0\,; \quad v_m^1(b) = 1.$$

The expressions for u_m and v_m are

$$u_m = r^{-N}\left[A_m(a)\, I_N(\beta_m r) + B_m(a)\, K_N(\beta_m r)\right], \tag{19}$$

$$v_m = r^{-N}\left[A_m(b)\, I_N(\beta_m r) + B_m(b)\, K_N(\beta_m r)\right], \tag{20}$$

where

$$A_m(r) = \frac{-r^N K_N(\beta_m r)}{\beta_m \Delta_m(r)},$$

$$B_m(r) = \frac{r^N I_N(\beta_m r)}{\beta_m \Delta_m(r)},$$

and $\Delta_m(r) = I_N(\beta_m r)\, K_{N+1}(\beta_m r) - I_{N+1}(\beta_m r)\, K_N(\beta_m r)$.

The corresponding Green's function is

$$\tilde{G}(r, \xi) = \frac{u_m(r)\, v_m(\xi)}{b^{2N+1}\, u_m(b)} \text{ for } r \leqslant \xi,$$

$$= \frac{u_m(\xi)\, v_m(r)}{b^{2N+1}\, u_m(b)} \text{ for } r > \xi.$$

Thus the solutions of the boundary value problem in f_m is given by [1],

$$f_m(r) = \int_a^b \xi^{2N+1} h_m(\xi)\, \tilde{G}_m(r, \xi)\, \mathrm{d}\xi. \tag{21}$$

Thus, boundary value problem for $V_m(r, t)$ and $\Phi_m(r, t)$ (equations (7) to (9)) now reduces to

$$-\alpha_m^2 C_{44} V_m + C_{66} \nabla_N^2 (V_m) + \alpha_m E_{14}\, \Phi_{m,r} - s V_{m,tt} = s G_{m,tt} - \alpha_m E_{14}\, F_{m,r} \tag{22}$$

$$\alpha_m E_{14}\left(V_{m,r} + \frac{2N+1}{r} V_m\right) + E_{11}\left(\Phi_{m,rr} + \frac{2N+1}{r}\Phi_{m,r}\right) - E_{33}\, \alpha_m^2\, \Phi_m = 0 \tag{23}$$

the boundary conditions being

$$\left.\begin{aligned} V_{m,r} - \frac{1}{r} V_m(r, t) &= 0 \\ \Phi_m(r, t) &= 0 \end{aligned}\right\} \text{ for } r = a, b. \tag{24}$$

4. Characteristic functions and frequency equation

Equations (22) to (24) are solved by adopting the procedure used in [3] as outlined below. Considering the homogeneous equations corresponding to (22) and assuming the expressions for V_m and Φ_m as

$$V_m = \sum_{n=1}^{\infty} V_{mn}(r) \exp(i w_{mn} t),$$

$$\Phi_m = \sum_{n=1}^{\infty} \Phi_{mn}(r) \exp(i w_{mn} t), \tag{25}$$

the equations and the boundary conditions are

$$(sw_{mn}^2 - \alpha_m^2 C_{44})\, V_{mn} + C_{66} \tilde{\nabla}^2 V_{mn} + E_{14}\alpha_m \Phi_{mn}^1 = 0,$$

$$\alpha_m E_{14}\left(V'_{mn} + \frac{2N+1}{r} V_{mn}\right) + E_{11}\left(\Phi_{mn}^{11} + \frac{2N+1}{r}\Phi'_{mn}\right)$$

$$- E_{33}\alpha_m^2 \Phi_{mn} = 0, \tag{26}$$

and

$$\left.\begin{aligned} &\Phi_{mn}(r) = 0 \\ &V_{mn}^1(r) - \frac{1}{r} V_{mn}(r) = 0 \end{aligned}\right\} \text{ for } r = a, b \tag{27}$$

where

$$\tilde{\nabla}^2 \equiv \left[\frac{\mathrm{d}^2}{\mathrm{d}r^2} + \frac{2N+1}{r}\frac{\mathrm{d}}{\mathrm{d}r} - \frac{2N+1}{r^2}\right](\ ).$$

Solutions of the system (26) are

$$V_{mn} = r^{-N} \sum_{i=1}^{2} [A_{mni} J_{N+1}(\delta_i r) + B_{mni} Y_{N+1}(\delta_i r)],$$

$$\Phi_{mn} = -\frac{C_{44}}{\alpha_m E_{14}} r^{-N} \sum_{i=1}^{2} d_{mni} [A_{mni} J_N(\delta_i r) + B_{mni} Y_N(\delta_i r)], \tag{28}$$

where $J_N(\ )$ and $Y_N(\ )$ are the Bessel functions of order N; A_{mni} and B_{mni} are arbitrary constants and

$$d_{mni} = \delta_i\, [\bar{C}_{66} + (\alpha_m/\delta_i)^2\, \Omega_{mn}].$$

δ_i^2 are the roots of the quadratic equation in δ^2

$$\bar{E}_{11}\bar{C}_{66}\delta^4 + (\bar{E}_{11}\Omega_{mn} + K_{14}^2 + \bar{C}_{66})\,\alpha_m^2\,\delta^2 + \Omega_{mn}\alpha_m^4 = 0, \tag{29}$$

where

$$\bar{E}_{11} = \frac{E_{11}}{E_{33}}\left(= \frac{\epsilon_{11}}{\epsilon_{33}}\right) \bar{C}_{66} = \frac{C_{66}}{C_{44}}\left(= \frac{c_{66}}{c_{44}}\right)$$

$$K_{14}^2 = \frac{E_{14}^2}{E_{33}C_{44}}\left(= \frac{e_{14}^2}{\epsilon_{33}c_{44}}\right) \text{ and } \Omega_{mn} = 1 - \frac{sw_{mn}^2}{C_{44}\alpha_m^2}\left(= 1 - \frac{\rho w_{mn}^2}{c_{44}\alpha_m^2}\right).$$

The frequency equation is therefore

$$\begin{vmatrix} d_{mn1}J_N(\delta_1 a) & d_{mn2}J_N(\delta_2 a) & d_{mn1}Y_N(\delta_1 a) & d_{mn2}Y_N(\delta_2 a) \\ d_{mn1}J_N(\delta_1 b) & d_{mn2}J_N(\delta_2 b) & d_{mn1}Y_N(\delta_1 b) & d_{mn2}Y_N(\delta_2 b) \\ \delta_1 J_{N+2}(\delta_1 a) & \delta_2 J_{N+2}(\delta_2 a) & \delta_1 Y_{N+2}(\delta_1 a) & \delta_2 Y_{N+2}(\delta_2 a) \\ \delta_1 J_{N+2}(\delta_1 b) & \delta_2 J_{N+2}(\delta_2 b) & \delta_1 Y_{N+2}(\delta_1 b) & \delta_2 Y_{N+2}(\delta_2 b) \end{vmatrix} = 0. \tag{30}$$

Table 1. Values of $L\Omega_{mn}/C_1$.

$n \backslash m$	1	2	3	4	5
1	3·1416	11·9958	13·9013	12·2191	14·6376
2	10·6907	18·1633	19·4744	16·1970	18·7393
3	17·3290	24·8005	25·7763	21·1739	23·1768
4	24·1962	31·7187	32·4874	27·0832	28·6763
5	31·2485	38·8034	39·4342	33·5338	34·8331

The characteristic functions V_{mn} satisfy the orthogonality conditions

$$\int_a^b r^{2N+1} V_{mn} V_{ml} \mathrm{d}r = 0 \text{ if } n \neq l,$$

$$\theta_{mn} = \int_a^b r^{2N+1} V_{mn}^2 \, \mathrm{d}r \neq 0. \tag{31}$$

5. Numerical results

The results in (29) are observed to be identical with those in (18) of [2], which deals with a similar problem in homogeneous medium. The frequency equation (30) is also similar to that given in (21) of [2] and the arguments of the corresponding Bessel functions are the same for both the frequency equations. Hence the numerical method of computing the roots of the frequency equation (30). in the present paper is exactly the same as in the special case for $N = 0$, which corresponds to the homogeneous medium. The roots of the frequency equation (30) may be represented as $L\Omega_{mn}/C_1$

where

$$C_1 = \frac{C_{44}}{S} (= c_{44}/\rho).$$

For the sake of completion the first few roots of the frequency equation (30) are presented in table 1 for β-quartz in the particular case corresponding to $N = 0$. These roots are reproduced from the present author's earlier paper, [2], which contains all the details of the numerical procedure adopted.

6. General solution of the boundary value problem

In terms of the characteristic functions the solutions of the system of equations (22) to (24) may be expressed as

$$V_m = \sum_{n=1}^{\infty} T_{mn}(t) \, V_{mn}(r),$$

$$\Phi_m = \sum_{n=1}^{\infty} T_{mn}(t)\, \Phi_{mn}(r). \tag{32}$$

The substitution of (32) in (22) to (24) gives the result

$$\ddot{T}_{mn} + w_{mn}^2 T_{mn} = q_{mn}(t), \tag{33}$$

where

$$q_{mn} = \alpha_m E_{14} a_{mn} P_m(t) - b_{mn} \ddot{P}_m,$$

and a_{mn} and b_{mn} are the generalized Fourier coefficients in the expansions of f_m' and g_m respectively as

$$g_m = \sum_{n=1}^{\infty} b_{mn} V_{mn}\ ; \quad f_m' = s \sum_{n=1}^{\infty} a_{mn} V_{mn}$$

Using the orthogonality conditions (31) the Fourier coefficients are given by

$$b_{mn} = \frac{1}{\theta_{mn}} \int_a^b r^{2N+1} V_{mn} g_m \mathrm{d}r\ ; \quad s a_{mn} = \frac{1}{\theta_{mn}} \int_a^b r^{2N+1} V_{mn} f_m^1 \,\mathrm{d}r.$$

The solution of (33) is found as

$$T_{mn} = M_{mn1} \cos(w_{mn} t) + M_{mn2} \sin(w_{mn} t) + \frac{1}{w_{mn}} \int_0^t$$

$$\times\, q_{mn}(x) \sin[w_{mn}(t-x)]\, \mathrm{d}x,$$

where M_{mn1} and M_{mn2} are arbitrary constants which may be determined from the initial conditions.

Finally the expressions for v and ϕ are obtained as

$$v = r^{-N} \sum_{m,n=1}^{\infty} \Big\{ (b_{mn} P_m(t) + M_{mn1} \cos(w_{mn} t) + M_{mn2} \sin(w_{mn} t)$$

$$+ \frac{1}{w_{mn}} \int_0^t q_{mn}(x) \sin(w_{mn}(t-x)\, \mathrm{d}x)$$

$$\times \sum_{i=1}^{2} [A_{mni} J_{N+1}(\delta_i r) + B_{mni} Y_{N+1}(\delta_i r)] \Big\} \cos(\alpha_m z), \tag{34}$$

$$\phi = r^{-N} \sum_{m,n=1}^{\infty} \Big\{ C_{mn} P_m(t) \sum_{i=1}^{2} [A_{mni} J_{N+1}(\delta_i r) + B_{mni} Y_{n+1}(\delta_i r)]$$

$$- \frac{c_{44}}{\alpha_m e_{14}} \Big[M_{mni} \cos(W_{mn} t) + M_{mn2} \sin(w_{mn} t)$$

$$+ \frac{1}{W_{mn}} \int_0^t q_{mn}(x) \sin(W_{mn}(t-x)\, \mathrm{d}x \Big]$$

$$\times \sum_{i=1}^{2} \mathrm{d}_{mni} [A_{mni} J_N(\delta_i r) + B_{mni} Y_N(\delta_i r)] \Big\} \sin(\alpha_m z), \tag{35}$$

where f_m is expressed as

$$f_m = \sum_{n=1}^{\infty} C_{mn} V_{mn} \text{ with } C_{mn} = \frac{1}{\theta_{mn}} \int_a^b r^{2N+1} V_{mn} f_m \mathrm{d}r.$$

7. Initial conditions

The initial conditions may be prescribed as

$$[v]_{t=0} = \bar{g}_1 (r, z) \; ; \quad [V,_t]_{t=0} = \bar{g}_2 (r, z). \tag{36}$$

Expanding the functions $\bar{g}_1$ and $\bar{g}_2$ in the form

$$\bar{g}_i (r,z) = \sum_{m=1}^{\infty} R_{mi} \cos (a_m z)$$

where

$$R_{mi} = \frac{2}{L} \int_0^L \bar{g}_i (r, z) \cos (a_m z) \, \mathrm{d}z, \; i = 1, 2. \tag{37}$$

and using the expressions (34) in (36) and (37), the relation involving the constants of integration M_{mni} is obtained as

$$R_{mi} = \sum_{n=1}^{\infty} (b_{mn} P_m (0) + w_{mn}^{i-1} M_{mni}) V_{mn}.$$

Application of the orthogonality condition (31) leads to the result

$$b_{mn} P_m (0) + w_{mn}^{i-1} M_{mni} = \frac{1}{\theta_{mn}} \int_a^b r^{2N+1} R_{mi} V_{mn} \mathrm{d}r \; i = 1, 2. \tag{38}$$

Hence the constants of integration M_{mni} in (34) and (35) are determined. From (1), (34) and (35) the distribution of stress and electric potential inside the shell is completely known.

8. Conclusions

Equations (34) and (35) give the expressions for the displacement v and the electric potential ϕ developed in the finite inhomogeneous cylindrical shell of (622) piezoelectric material wherein the inhomogeneity is restricted to the variations of all the physical constants of the medium as r^{2N} and the shell is subjected to a time-dependent shearing stress.

Equations (22) and (23) of [3] give the displacement and the electric potential of a similar problem but with time-dependent electrical boundary condition. The following two observations are made for comparison. Firstly, if $p(z,t)$ is made

zero the expressions (34) and (35) of the present solution and those of [3], given in (22) and (23) are identical. This particular case is the solution of a trivial problem under nonzero initial conditions and when the entire boundary is free from both traction and applied electrical potential.

Secondly the electric potential ϕ appearing in (23) of [3] can be considered as a superposition of two outcomes, one due to the applied electric potential, represented by $P(z,t)$ and the other due to the potential developed in the trivial problem, free of both the applied tractions and electric potential. Similarly the torsional oscillations described by v in (34) can be viewed as a combination of two waves ; one due to the applied traction given by $p(z,t)$ and the other due to the torsional wave with time-dependent electrical potential, obtained in (22) of [3]. By a proper choice of $P_m(t)$, the phase difference may be controlled to yield a wave form with periodic pulses. Finally as observed earlier, the frequency equation (30) is very much similar to the frequency equation (18) in [2] wherein the piezoelectric medium is homogeneous. Hence the type of inhomogeneity considered in the present investigation leads to a very simple generalization of a similar problem in homogeneous medium. This also follows from the observation that the arguments of the Bessel functions appearing in the frequency equation (30) are closely related to the roots of (29) which remains unchanged whether the medium is homogeneous or inhomogeneous of the type investigated and also whether the shell is subjected to time-dependent electrical or mechanical boundary condition.

Acknowledgements

The author is grateful to Prof. Venkataraman of PSG College of Technology for his kind suggestions and discussions in the preparation of the paper. The author is also grateful to the referee for his comments and suggestions which led to the revision of the paper.

References

[1] Churchill R V 1958 *Operational Mathematics* second edition, International Student edtion, (———— : McGraw-Hill) p. 262
[2] Paul H S and Sarma K V *Proc. Indian Natl. Sci. Acad.* **A43** 169
[3] Sarma K V 1980 *Int. J. Eng. Sci.* **18** 449
[4] Tranter C J 1968 *Bessel functions with some physical applications* (The English Universities Press Ltd.) p. 58

Proc. Indian Acad. Sci. (Math. Sci.), Vol. 91, Number 3, November 1982, pp. 211–216.

Slow motion of a micropolar fluid through a porous sphere bounded by a solid sphere

B S BHATT

Department of Mathematics, The University of the West Indies, St. Augustine, Trinidad, West Indies

MS received 22 February 1982 ; revised 5 August 1982

Abstract. The paper examines the slow motion of a micropolar fluid produced by the relative motion of a solid sphere to an inside porous sphere. The result extends the Cunningham's problem to micropolar fluid when the inner sphere is porous with prescribed radial suction/injection velocity at the surface of the sphere. The result can also be taken as an extension of the work of Ramkissoon and Majumdar when the fluid is bounded at a radius $r = b$ $(b > a)$ but the solid sphere is replaced by a porous sphere. The force experienced by the inner sphere has been calculated and particular cases of interest have been deduced.

Keywords. Micropolar fluid ; Stokes' flow ; suction/injection ; relative motion ; drag formula.

1. Introduction

A lot of work has been carried out in micropolar fluid after the introduction of the theory of the fluids given by Eringen [1]. Here we shall concentrate on the work done by the various authors in flow past spheres using Eringen's fluid model in different cases. Ramkissoon and Majumdar [2] have discussed the creeping flow of a micropolar fluid past a sphere. The same problem was also looked by Lakshmana Rao and Bhujanga Rao [3]. A more general result for Stokes' flow problem of micropolar fluids in the case of axisymmetric bodies has been given in another paper by Ramkissoon and Majumdar [4]. Expression for the drag has been obtained and applied for the case of a sphere.

Slow motion due to the rotation of a sphere in micropolar fluid has also been the interest of study. Lakshmana Rao *et al* [5] discussed the problem using Stokes' approximation, whereas more general results have been obtained by Ramkissoon [6] for the same problem.

For concentric spheres only the secondary flow induced by the rotation of spheres has been discussed by Bhatnagar [7].

Here an attempt has been made to extend Cunningham's result [8] for micropolar fluid by replacing the inner solid sphere by a porous sphere with prescribed

radial boundary suction/injection velocity $m_0 U \cos\theta$ ($m_0 \gtrless 0$ for suction/injection). The force experienced by the inner porous sphere has been calculated and discussed in some particular cases.

2. Formulation of the problem

We consider a porous sphere of radius a placed at the centre of a micropolar fluid. Another solid sphere of radius $b > a$ bounds the fluid. The solid sphere has a relative velocity U to the inner porous sphere in the negative direction of z-axis.

The boundary conditions to be satisfied are :

(i) The velocity components be finite at the origin,

(ii) $q_{r1} = m_0 U \cos\theta$; $q_{\theta 1} = 0$ at $r = a$, (1)

(iii) $\nu_2 = 0$; $q_{r2} = m_0 U \cos\theta$ $q_{\theta 2} = 0$; at $r = b$ (2)

(iv) $\psi_2 = \frac{1}{2} r^2 U \sin^2\theta$ $\partial\psi_2/\partial r = rU \sin^2\theta$; at $r = b$ (3)

where q_{ri}, $q_{\theta i}$, ($i = 1, 2$) are the velocity components of the fluid inside and outside the porous sphere, ψ_2, ν_2 are the stream function and microrotation for outside flow field of porous sphere, *i.e.*, for $a < r < b$.

Flow equations are exactly the same as that of [2]. So we can write in non-dimensional form, the flow equations for $0 < r < 1$ and $1 < r < \sigma$ ($= b/a$) by taking $i - 1, 2$ (using [2]).

The stream function ψ_i and micro-rotation ν_i satisfy

$$L_{-1}^2 (L_{-1} - \lambda_1^2)\psi_i = 0, \tag{4}$$

and

$$\nu_i = \frac{1}{2r\sin\theta}\left(L_{-1}\psi i + \frac{N_2 + N_3}{N_3^2} L_{-1}^2 \psi i\right) \tag{5}$$

where

$\lambda_1^2 = N_3(2N_2 + N_3)/(N_2 + N_3)$ and L_{-1} is the axisymmetric operator

$$L_{-1} = \frac{\partial^2}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2}{\partial\theta^2} - \frac{\cot\theta}{r^2}\frac{\partial}{\partial\theta}.$$

The boundary conditions in non-dimensional form can be written as

(a) q_{r1} and $q_{\theta 1}$ be finite at $r = 0$, (6)

$$\text{(b)}\quad \left.\begin{aligned} q_{r1} &= m_0 \cos\theta \\ q_{\theta 1} &= 0 \end{aligned}\right\} \text{ at } r = 1, \tag{7}$$

$$\text{(c)}\quad \left.\begin{aligned} \nu_2 &= 0 \\ q_{r2} &= m_0 \cos\theta \\ q_{\theta 2} &= 0 \end{aligned}\right\} \text{ at } r = 1, \tag{8}$$

$$\text{(d)}\quad \left.\begin{aligned} \psi_2 &= \tfrac{1}{2} r^2 \sin^2\theta, \\ \frac{\partial\psi_2}{\partial r} &= r\sin^2\theta \end{aligned}\right\} \text{ at } r = \sigma. \tag{9}$$

3. Solution

After some straightforward manipulations the solution can be obtained as

$$\psi_1 = -\frac{m_0}{2} r^2 \sin^2\theta\,(2 - r^2), \tag{10}$$

$$v_1 = \frac{5}{2} m_0 r \sin\theta, \tag{11}$$

$$\psi_2 = \frac{1}{2}\sin^2\theta\left[\frac{A_2}{r} + B_2 r + C_2 r^2 + D_2 r^4 + E_2\left(\lambda_1 + \frac{1}{r}\right)e^{-\lambda_1 r}\right], \tag{12}$$

$$v_2 = \frac{1}{2r}\sin\theta\left[E_2\lambda_1^2\frac{N_2+N_3}{N_3}\left(\lambda_1 + \frac{1}{r}\right)e^{-\lambda_1 r} - \frac{B_2}{r} + 5D_2 r^2\right] \tag{13}$$

where

$$\begin{aligned} A_2 &= -\frac{2}{3}m_0 + [6k_1(\phi_1\phi_4 - \phi_2\phi_3)]^{-1}[(3+m_0)\sigma\{(\lambda_1^2 e^{-\lambda_1} - 2k_1)\phi_4 \\ &\quad + (\lambda_1^2 e^{-\lambda_1} + k_1)5\phi_3 - 9k_1\phi_3 - 3e^{-\lambda_1}(2 + 2\lambda_1 + \lambda_1^2)(\phi_4 + 5\phi_3)\} \\ &\quad - \left\{\sigma^3 + \frac{m_0}{3}(2+\sigma^3)\right\}\{[\lambda_1^2 e^{-\lambda_1} - 2k_1]\phi_2 + [\lambda_1^2 e^{-\lambda_1} + k_1]5\phi_1 \\ &\quad - 9k_1\phi_1 - 3e^{-\lambda_1}(2 + 2\lambda_1 + \lambda_1^2)(\phi_2 + 5\phi_1)\}), \end{aligned} \tag{14}$$

$$B_2 = (\phi_1\phi_4 - \phi_2\phi_3)^{-1}\left[(3+m_0)\sigma\phi_4 - \left\{\sigma^3 + \frac{m_0}{3}(2+\sigma^3)\right\}\phi_2\right], \tag{15}$$

$$\begin{aligned} C_2 &= -\frac{m_0}{3} + [3k_1(\phi_1\phi_4 - \phi_2\phi_3)]^{-1}\Big[(3+m_0)\sigma\{(\lambda_1^2 e^{-\lambda_1} - 2k_1)\phi_4 \\ &\quad + [\lambda_1^2 e^{-\lambda_1} + k_1]5\phi_3\} - \left\{\sigma^3 + \frac{m_0}{3}(2+\sigma^3)\right\}\{[\lambda_1^2 e^{-\lambda_1} - 2k_1]\phi_2 \\ &\quad + [\lambda_1^2 e^{-\lambda_1} + k_1]5\phi_1\}\Big], \end{aligned} \tag{16}$$

$$D_2 = (\phi_1\phi_4 - \phi_2\phi_3)^{-1}\left[\left(\sigma^3 + \frac{m_0}{3}(2+\sigma^3)\right)\phi_1 - (3+m_0)\sigma\phi_3\right], \tag{17}$$

$$\begin{aligned} E_2 &= [k_1(\phi_1\phi_4 - \phi_2\phi_3)]^{-1}\Big[(3+m_0)\sigma(\phi_4 + 5\phi_3) - \left(\sigma^3 + \frac{m_0}{3}(2+\sigma^3)\right) \\ &\quad \times(\phi_2 + 5\phi_1)\Big], \end{aligned} \tag{18}$$

with

$$\phi_1 = 2(1-\sigma) - \frac{\lambda_1^2}{k_1}(e^{-\lambda_1\sigma} - \sigma e^{-\lambda_1}),$$

$$\phi_2 = 5\left[\sigma(\sigma^2 - 1) + \frac{\lambda_1^2}{k_1}(e^{-\lambda_1\sigma} - \sigma e^{-\lambda_1})\right],$$

$$\phi_3 = -\frac{1}{3}(1-\sigma)^2(1+2\sigma) + \frac{\lambda_1}{k_1}\left\{(\sigma - \lambda_1 - 1)e^{-\lambda_1\sigma} + \frac{\lambda_1}{3}(2+\sigma^3)e^{-\lambda_1}\right\},$$

$$\phi_4 = \frac{1}{3}(1-\sigma)^2 (2 + 4\sigma + 6\sigma^2 + 3\sigma^3) - \frac{5\lambda_1}{k_1}\left\{(\sigma - \lambda_1 - 1) \times e^{-\lambda_1\sigma} + \frac{\lambda_1}{3}(2+\sigma^3)\, e^{-\lambda_1}\right\},$$

$$k_1 = \lambda_1^2 \frac{N_2 + N_3}{N_3}(\lambda_1 + 1)\, e^{-\lambda_1}.$$

Force on the porous sphere (inner) may be obtained as

$$D = 2\pi (2N_2 + N_3)B_2$$

$$= 2\pi (2N_2 + N_3) \frac{(3 + m_0)\sigma\phi_4 - \left\{\sigma^3 + \frac{m_0}{3}(2 + \sigma^3)\right\}\phi_2}{\phi_1\phi_4 - \phi_2\phi_3}. \tag{19}$$

For $m_0 = 0$, (19) gives the results for the solid sphere bounded by a solid sphere and we get

$$D = 2\pi (2N_2 + N_3) \frac{3\sigma\phi_4 - \sigma^3\phi_2}{\phi_1\phi_4 - \phi_2\phi_3}. \tag{20}$$

which is the corresponding result to that of Cunningham [8] in micropolar fluid. As $\sigma \to \infty$, (19) reduces to

$$D = -\frac{2\pi (2N_2 + N_3)(N_2 + N_3)(1 + \lambda_1)(3 + m_0)}{N_3 + 2N_2 + 2\lambda_1 N_3 + 2\lambda_1 N_2} \tag{21}$$

which is the drag on a porous sphere bounded by an infinite mass of micropolar fluid.

For $m_0 = 0$ in (21) we get the drag on a solid sphere as obtained by Ramkissoon and Majumdar [2] as,

$$D = \frac{-6\pi (2N_2 + N_3)(N_2 + N_3)(1 + \lambda_1)}{N_3 + 2N_2 + 2\lambda_1 N_3 + 2\lambda_1 N_2} \tag{22}$$

Taking $N_3 = 0$ *i.e.* $\lambda_1 = 0$ in (19)–(22), we get the corresponding results for Newtonian fluid as (in dimensional form)

$$D = \frac{-6\pi\mu Ua\left[(1 - \lambda^5) + \frac{m_0}{3}(1 + 5\lambda^3 - 6\lambda^5)\right]}{1 - 9\lambda/4 + 5\lambda^3/2 - 9\lambda^5/4 + \lambda^6}, \tag{23}$$

$$D = \frac{-6\pi\mu Ua(1 - \lambda^5)}{1 - 9\lambda/4 + 5\lambda^3/2 - 9\lambda^5/4 + \lambda^6} \tag{24}$$

$$D = -2\pi\mu Ua(3 + m_0) \tag{25}$$

$$D = -6\pi\mu Ua, \tag{26}$$

with $\lambda = \frac{a}{b}$ respectively.

(24) agrees with the result of Cunningham [8], as also obtained by Happel and Brenner [9]. (25) and (26) are the classical drag formulas for porous and non-porous spheres placed in an infinite mass of Newtonian fluid of viscosity μ.

4. Numerical discussions

Figure 1 depicts the behaviour of D/D_c [where $D_c = -6\pi N_2$ and D is given by (19)]. As one would expect, when the fluid in and around the porous sphere is bounded by a solid sphere the permeability m_0 will have to be restricted in order that a flow is possible. That is for given σ, λ_1 and N_3/N_2 ; m_0 will have to be less than some quantity. Similarly if we fix m_0, λ_1 and N_3/N_2 we shall get the upper bound for σ. The same has been observed during calculations. The similar situation one would get for Newtonian fluid also. The required condition connecting σ, λ_1, N_3/N_2 and m_0 can be obtained by the requirement that D/D_c should always be positive. The curves in the figure are similar to what

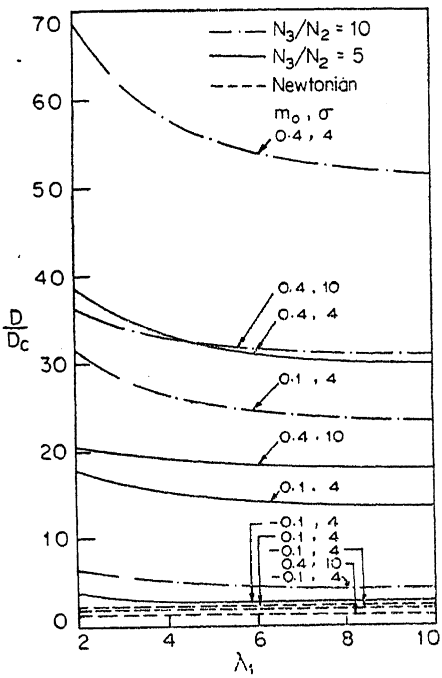


Figure 1. The Drag.

Table 1. The drag for Newtonian case.

m_0	σ	D
0·1	4	2·1627
—0·1	4	2·0127
0·4	4	2·1777
—0·4	4	1·7977
0·4	10	1·4591
—0·4	10	1·1133

Ramkissoon and Majumdar [2] have obtained. For the same σ increase of m_0 increases D/D_c (for $m_0 > 0$) and increase of m_0 will decrease D/D_c (for $m_0 < 0$) whereas for same m_0, D/D_c decreases with the increase of σ in the case of Newtonian fluid (see table 1). The same behaviour continues for micropolar fluid though the values in micropolar case are much higher than the Newtonian case. This again is in agreement with Ramkissoon's and Majumdar's [2] findings.

In the case of a porous sphere placed in an infinite mass of micropolar fluid, (21) shows that Ramkissoon's and Majumdar's [2] drag formula is multiplied by $(3 + m_0)$. Thus one can easily get the behaviour of D/D_c. Since $m_0 \ll 1$, therefore there will be a little shift in the curves obtained in [2] above and below for $m_0 >$ or < 0 respectively.

Acknowledgements

The author is thankful to the referee for valuable suggestions.

References

[1] Eringen A C 1964 *Int. J. Eng. Sci.* **2** 205
[2] Ramkissoon H and Majumdar S R 1975 *Lett. Appl. Eng. Sci.* **3** 133
[3] Lakshmana Rao S K and Bhujanga Rao P 1970 *J. Eng. Math.* **4** 209
[4] Ramkissoon H and Majumdar S R 1976 *Phys. Fluids* **19** 16
[5] Lakshmana Rao S K, Ramacharyulu N Ch. P and Bhujanga Rao P 1969 *Int. J. Engg. Sci.* **7** 905
[6] Ramkissoon H 1977 *Appl. Sci. Res.* **33** 243
[7] Bhatnagar K S 1969 *Rheol. Acta* **8** 44
[8] Cunningham E 1910 *Proc. R. Soc.* **A83** 357
[9] Happel J and Brenner H 1973 *Low Reynolds number hydrodynamics* (Noordhoff Int. Pub.) p. 132

Proc. Indian Acad. Sci. (Math. Sci.), Vol. 91, Number 3, November 1982, pp. 217-221.

Zero-free regions of derivatives of Riemann zeta function

D P VERMA and A KAUR
Department of Mathematics, Patna University, Patna 800 005, India

MS received 20 October 1981

Abstract. Zero-free regions of the kth derivative of the Riemann zeta function $\zeta^{(k)}(s)$ are investigated. It is proved that for $k \geq 3$, $\zeta^{(k)}(s)$ has no zero in the region $\mathrm{Re}\, s \geq (1 \cdot 1358826 \ldots) k + 2$. This result is an improvement upon the hitherto known zero-free region $\mathrm{Re}\, s \geq (7/4) k + 2$ on the right of the imaginary axis. The known zero-free region on the left of the imaginary axis is also improved by proving that $\zeta^{(k)}(s)$ may have at the most a finite number of non-real zeros on the left of the imaginary axis which are confined to a semicircle of finite radius r_k centred at the origin.

Keywords. Riemann zeta function ; zero-free region.

1. Introduction

There has been great interest and activity in locating the zero-free regions [4] of Riemann zeta function $\zeta(s)$. Recently some work has been done to study the behaviour of the derivatives of $\zeta(s)$ from this angle (1, 3, 5-7). Spira [5] has obtained the following results in this direction.

(a) If $k \geq 3$ and $\sigma \geq \frac{7}{4} k + 2$, then $\zeta^{(k)}(s) \neq 0$. (1)

(b) For each integer $k > 1$ and each $\epsilon > 0$ there exists an r_k such that $\zeta^{(k)}(s) \neq 0$ in the region defined by $\sigma < -\epsilon$, $|t| > \epsilon$, $|s| > r_k$. (2)

The first of these results deals with the region on the right of the imaginary axis and states that the entire complex plane on the right of $\mathrm{Re}\, s = (7/4) k + 2$ is free from zeros of $\zeta^{(k)}(s)$ and the second result states that on the left of the imaginary axis, the region $\mathrm{Re}\, s < -\epsilon$, but for a semicircle of finite radius r_k, is free from non-real zeros of $\zeta^{(k)}(s)$. He also remarks that one cannot hope to prove that the left half plane is free from non-real zeros of $\zeta^{(k)}(s)$.

We improve both the results (1) and (2). We add a vertical strip $(1 \cdot 1358826 \ldots) k + 2 \leq \mathrm{Re}\, s < (7/4) k + 2$ to the zero-free region (1) of $\zeta^{(k)}(s)$ on the right of the imaginary axis. On the left of the imaginary axis we extend the zero-free region (2) to include the strip $-\epsilon < \mathrm{Re}\, s \leq 0$ outside the semicircle. Thus in the next two sections, we prove the following new theorems.

Theorem (*A*) : If $k \geq 3$ and Re $s \geq ak + 2$ where

$$a = \frac{\log \dfrac{\log 3}{\log 2}}{\log \dfrac{3}{2}} = 1{\cdot}1358826\ldots$$

then $\zeta^{(k)}(s) \neq 0$. (3)

Theorem (*B*) : For each integer $k \geq 0$ there exists an r_k such that $\zeta^{(k)}(s)$ has no non-real zero in the region Re $s \leq 0$ except atmost a finite number in the semi-circle $|s| \leq r_k$. (4)

2. Zero-free region on the right : Proof of theorem (A)

For Re $s \geq a > 1$, the Dirichlet series for $\zeta(s)$ is uniformly convergent. Differentiating this series k times,

$$\zeta^{(k)}(s) = (-1)^k \sum_{n=2}^{\infty} \frac{\log^k n}{n^s}. \tag{5}$$

Thus

$$|\zeta^{(k)}(s)| \geq \frac{\log^k 2}{2^\sigma} - \sum_{n=3}^{\infty} \frac{\log^k n}{n^\sigma} \tag{6}$$

For $\sigma \geq k$, the monotonic character of $\log^k x/x^\sigma$ for $x > e$ gives

$$\sum_{n=5}^{\infty} \frac{\log^k n}{n^\sigma} \leq \int_4^\infty \frac{\log^k x}{x^\sigma}\, dx.$$

Integrating by parts several times, we get

$$\int_4^\infty \frac{\log^k x}{x^\sigma}\, dx = \frac{4^{1-\sigma}}{\sigma - 1} \sum_{r=0}^{k} \frac{k\,!}{(k-r)\,!} \cdot \frac{\log^{k-r} 4}{(\sigma-1)^r}$$

$$\leq \frac{4^{1-\sigma} \log^k 4}{\sigma - 1} \left\{ 1 + \sum_{r=1}^{k} \frac{k(k-1)^{r-1}}{(\sigma-1)^r \log^r 4} \right\}$$

$$< \frac{4^{1-\sigma} \log^k 4}{\sigma - 1} \left\{ 1 + \frac{k}{(\sigma-1)\log 4} \cdot \frac{1}{1 - \dfrac{k-1}{(\sigma-1)\log 4}} \right\}$$

Hence

$$|\zeta^{(k)}(s)| > \frac{\log^k 2}{2^\sigma} - \frac{\log^k 3}{3^\sigma} - \frac{\log^k 4}{4^\sigma} - \frac{4^{1-\sigma} \log^k 4}{\sigma - 1} \times \left\{ 1 + \frac{k}{(\sigma-1)\log 4 - k + 1} \right\}$$

$$= \frac{\log^k 3}{3^\sigma}\left(\left(\frac{3}{2}\right)^\sigma \left(\frac{\log 2}{\log 3}\right)^k - 1\right) - \frac{\log^k 4}{4^\sigma}\left[1 + \frac{4}{\sigma - 1} \times \left\{1 + \frac{k}{(\sigma-1)\log 4 - k + 1}\right\}\right] \tag{7}$$

Taking $\sigma \geq ak + 2$ where

$$a = \frac{\log \dfrac{\log 3}{\log 2}}{\log \dfrac{3}{2}} = 1 \cdot 1358826\ldots \tag{8}$$

it is easily seen that

$$\left(\frac{3}{2}\right)^\sigma \left(\frac{\log 2}{\log 3}\right)^k - 1 \geq \frac{5}{4}$$

and that

$$1 + \frac{4}{\sigma - 1}\left(1 + \frac{k}{(\sigma - 1)\log 4 - k + 1}\right) \leq 1 + \frac{4}{ak+1} \times \left(1 + \frac{k}{(ak+1)\log 4 - k + 1}\right)$$

$$= 1 + \frac{4}{ak+1} \cdot \frac{(ak+1)\log 4 + 1}{(ak+1)\log 4 - k + 1}$$

Thus (7) gives

$$|\zeta^{(k)}(s)| > \frac{5\log^k 3}{4 \cdot 3^\sigma} - \frac{\log^k 4}{4^\sigma}\left(1 + \frac{4}{ak+1} \cdot \frac{(ak+1)\log 4 + 1}{(ak+1)\log 4 - k + 1}\right)$$

$$\geq \frac{\log^k 4}{4^\sigma}\left\{\frac{5}{4}\left(\frac{4}{3}\right)^{ak+2}\left(\frac{\log 3}{\log 4}\right)^k - 1 - \frac{4}{ak+1} \cdot \frac{(ak+1)\log 4 + 1}{(ak+1)\log 4 - k + 1}\right\}$$

$$= \frac{\log^k 4}{4^\sigma}\left(\frac{20 \cdot 2^{(a-1)k}}{9} - 1 - \frac{4}{ak+1} \cdot \frac{(ak+1)\log 4 + 1}{(ak+1)\log 4 - k + 1}\right).$$

The expression in the large brackets increases monotonically with k, hence for $k \geq 3$, we have

$$|\zeta^{(k)}(s)| > \frac{\log^k 4}{4^\sigma}\left(\frac{20 \cdot 2^{3(a-1)}}{9} - 1 - \frac{4}{3a+1} \cdot \frac{(3a+1)\log 4 + 1}{(3a+1)\log 4 - 2}\right)$$

$$= \frac{\log^k 4}{4^\sigma}(0 \cdot 377) > 0.$$

This proves the theorem.

3. Zero-free region on the left : Proof of theorem (B)

We first prove that if $\zeta^{(k)}(s)$ has only a finite number of non-real zeros in $\operatorname{Re} s \leq 0$ then $\zeta^{(k+1)}(s)$ has the same property. (9)

$\zeta^{(k)}(s)$ is real on the real axis, so its non-real zeros are in conjugate pairs. Let these be $p_n \pm iq_n$, where $q_{n+1} \geq q_n > 0$ for all n, and let the real zeros be $-a_n$.

Berndt [1] proved that if $N_k(T)$ be the number of zeros of $\zeta^{(k)}(s)$ in $0 < t < T$ then

$$N(T) = N_k(T) + \frac{T \log 2}{2\pi} + O(\log T),$$

where [8]

$$N(T) = \frac{T}{2\pi} \log \frac{T}{2\pi} - \frac{T}{2\pi} + O(\log T).$$

Hence the infinite product

$$\Pi\left[\left(1 - \frac{s}{p_n + iq_n}\right)\left(1 - \frac{s}{p_n - iq_n}\right)\right]$$

is convergent. The only pole of $\zeta^{(k)}(s)$ is of order $k+1$ at $s = 1$, and $(s-1)^{k+1} \zeta^{(k)}(s)$ is an integral function of order 1, so

$$\zeta^{(k)}(s) = \frac{e^{b+cs}}{(s-1)^{k+1}} \Pi\left[\left(1 - \frac{s}{p_n + iq_n}\right)\left(1 - \frac{s}{p_n - iq_n}\right)\right]$$
$$\times \Pi\left(1 + \frac{s}{a_n}\right) e^{-s/a_n}.$$

Taking logarithms and differentiating,

$$\frac{\zeta^{(k+1)}(s)}{\zeta^{(k)}(s)} = c - \frac{k+1}{s-1} + \sum \left(\frac{1}{s - p_n - iq_n} + \frac{1}{s - p_n + iq_n}\right)$$
$$+ \sum \left(\frac{1}{s + a_n} - \frac{1}{a_n}\right),$$

so that

$$\operatorname{Re} \frac{\zeta^{(k+1)}(s)}{\zeta^{(k)}(s)} = O(1) + \sum \left(\frac{\sigma - p_n}{|s - p_n - iq_n|^2} + \frac{\sigma - p_n}{|s - p_n + iq_n|^2}\right)$$
$$+ \sum \left(\frac{\sigma + a_n}{|s + a_n|^2} - \frac{1}{a_n}\right).$$

We assumed that $\zeta^{(k)}(s)$ has atmost a finite number of non-real zeros in the half plane $\operatorname{Re} s \leq 0$, so the corresponding terms in the first summation give a finite sum while its other terms are all negative for $\operatorname{Re} s \leq 0$. Hence

$$\operatorname{Re} \frac{\zeta^{(k+1)}(s)}{\zeta^{(k)}(s)} < O(1) + \sum \left(\frac{\sigma + a_n}{|s + a_n|^2} - \frac{1}{a_n}\right)$$
$$= O(1) - \sum \frac{|s|^2}{a_n |s + a_n|^2} - \sum \frac{\sigma}{|s + a_n|^2}. \tag{10}$$

By Spira [5, 6] there is a constant A_k such that $\zeta^{(k)}(s)$ has no non-real zeros for $\operatorname{Re} s < -A_k$. Spira [6] also proves that

$$a_n = 2n + O(1).$$

Hence for $-A_k \leq \sigma \leq 0$ and t large, the second summation in (10) is of $O(1)$, and

$$\sum \frac{|s|^2}{a_n |s + a_n|^2} > \sum_{a_n < |s|/2} \frac{|s|^2}{a_n |s + a_n|^2}$$

$$> \sum_{a_n < |s|/2} \frac{4}{9a_n}$$

$$= O(1) + \frac{4}{9} \log |s|$$

so that

$$\operatorname{Re} \frac{\zeta^{(k+1)}(s)}{\zeta^{(k)}(s)} < O(1) - \frac{4}{9} \log |s|.$$

Thus $\quad \zeta^{(k+1)}(s) \neq 0$ for t large, say $t > t_k$. (11)

By theorem (*B*), $\zeta^{(k+1)}(s)$ has no non-real zeros for $\sigma < -A_k$.

Hence $\zeta^{(k+1)}(s)$ has the whole of the half plane $\sigma \leq 0$ free of non-real zeros except possibly for a finite part given by

$$-A_k \leq \sigma \leq 0, \quad 0 < \epsilon \leq t \leq t_k.$$

As there is no limit point of zeros of $\zeta^{(k+1)}(s)$, so it can have atmost a finite number of zeros in this finite bounded region. This proves (9).

Finally we prove Theorem (B) with the help of (9).

Let M be the set of natural numbers k for which $\zeta^{(k)}(s)$ has no non-real zeros in the half plane $\sigma \leq 0$ but for a finite number.

We know that $\zeta(s)$ has no non-real zeros in $\sigma \leq 0$. Hence $0 \in M$. (12)

By (9).

$$k \in M \Rightarrow k + 1 \in M. \tag{13}$$

By the law of induction [2], we conclude from (12) and (13) that M is the set of all natural numbers.

Let the non-real zero of largest modulus in $\sigma \leq 0$ be s_k and let r_k be any real number greater than or equal to $|s_k|$. Then theorem (B) follows.

References

[1] Berndt B 1970 *J. London Math. Soc.* **2** 577

[2] Eves H and Newsom C V 1965 *An introduction to the foundations and fundamental concepts of mathematics* (New York : Holt, Rinchart and Winston), Rev. Ed., p. 209

[3] Levinson N and Montgomery H L 1974 *Acta Math.* **133** 49

[4] Ramachandra K 1979 in *Riemann zeta function* (Madras : Ramanujan Inst. for Adv. Study in Maths) p. 6

[5] Spira R 1965 *J. London Math. Soc.* **40** 677

[6] Spira R 1970 *Proc. Am. Math. Soc.* **26** 246

[7] Spira R 1973 *Ill. J. Math.* **17** 147

[8] von Mangoldt H 1905 *Math. Ann.* **60** 1

Proc. Indian Acad. Sci. (Math. Sci.), Vol. 91, Number 3, November 1982, pp. 223–234.

Halphen Puiseux inequalities in the precessional motion of a rolling missile

P C RATH and JAI PAL
Institute of Armament Technology, Girinagar, Pune 411 025, India

MS received 5 February 1982

Abstract. In the theory of precessional motion of the rolling missile, [5] observed that the aspidal angle is always bounded below and above by $\pi/2$ and π respectively, provided the roots of the libration polynomial satisfy certain launching conditions. In this paper, we have obtained the same results without these conditions. We have also proved that the bounds are sharp, so far as the purely retrograde and direct motions are concerned.

Keywords. Precessional limits ; rolling missile
missile.

1. Introduction

It is well known that the apsidal angle ϕ of a spherical pendulum is such that it always satisfies the following inequalities

$$\pi/2 < \Phi < \pi \tag{1}$$

for all conditions of motion and(1) is known as the Halphen Puiseux inequalities. The same angle Φ for a common top also satisfies (1) for purely retrograde motion and for direct motion, although the first inequality (Halphen's) is satisfied, the second one (Puiseux) is satisfied only asymptotically. For this reference may be made to Kohn [3] and Diaz and Metcalf [2]. The precessional motion of a rolling missile which is modelled like a common top (The Lock-Fowler missile for example, *cf* Rath and Namboodiri [4], [5], [6] has a lot of qualitative agreements with its simulator, *i.e.* the common top). Qualitative differences between the missile and its simulator do exist because the impressed forces of the missile are aerodynamic in nature. In case of grape-vine motions of a missile precession does not always have the same sign as spin [5]. It is therefore obvious that the apsidal angle Φ of a rolling missile of the Lock-Fowler type will not have the same limits as its simulator, for all conditions of motion. We have already proved * that the apsidal angle of a Lock-Fowler missile satisfies Halphen's

* The same result is also proved here by a different method (See Rath and Pal).

inequality for purely direct and retrograde motions [7]. Here we shall prove that in case of direct and retrograde motion (of the missile) Φ satisfies the Puiseux inequality asymptotically and π and $\pi/2$ are the least upper bound and the greatest lower bound respectively of Φ for purely direct and retrograde motions only.

2. Basic equations and the apsidal angle

With notations same as in [5], the equations characterising the nutational and precessional motion of the missile may be written as follows

$$(\mathrm{d}z/\mathrm{d}t)^2 = H(z) \tag{2}$$

and

$$(\mathrm{d}\phi/\mathrm{d}t) = \Omega(\lambda - z)/2z(1 - z) \tag{3}$$

where

$$z = 2^{-1}(1 + \cos\delta), \tag{4}$$

$$\lambda = (F + \Omega)/2\Omega, \tag{5}$$

$$\Omega = AN/B, \tag{6}$$

$$\left.\begin{aligned} H^{**}(z) &= z(1 - z)\,[E_1 - \alpha(2z - 1) - \beta(2z - 1)^2] - \Omega^2(\lambda - z)^2 \\ &= 4\beta\left[\prod_{i=1}^{4}(z - z_i)\right], \\ \text{with } E_1 &= E + \alpha + \beta. \end{aligned}\right\} \tag{7}$$

Here ϕ and δ are the angles of precession and nutation respectively, E and F are constants representing energy and angular momentum. [(See [5], equations (2·1) and (2·2)]. A and B are the axial and transverse moments of inertia and N is the axial spin of the missile. The parameters α, β and s are given by

$$\left.\begin{aligned} \alpha &= (\Omega^2/2s)(1 - 4qs), \\ \beta &= q\Omega^2, \\ s &= A^2N^2/4B\mu(0), \end{aligned}\right\} \tag{8}$$

where $\mu(0)$ is a certain aerodynamic parameter with reference to linear angular motion of the missile. It may be noted that for real motions z_i $(i = 1, 2, \ldots, 4)$, the zeros of the polynomial $H(z)$ are real and satisfy the following inequalities

$$z_1 \leqslant 0 \leqslant z_2 < z < z_3 \leqslant 1 \leqslant z_4 \tag{9}$$

** This equation is the correct version of (2.10) of Rath and Namboodiri [5] for it can be seen that their (2.6) contains an algebraic error, the corrected form of which is

$$\delta'^2\sin^2\delta + (F - \Omega\cos\delta)^2 + \frac{\Omega^2\sin^2\delta}{2s}\,[(1 - 4qs)\cos\delta + 2qs\cos^2\delta]_0^{\delta} = E\sin^2\delta \tag{7}$$

Defining certain symmetric function of the roots as

$$L = \left[- \prod_{i=1}^{4} z_i \right]^{1/2} \tag{10}$$

and

$$L' = \left[- \prod_{i=1}^{4} (1 - z_i) \right]^{1/2} \tag{11}$$

It can be shown that (see Rath and Namboodiri [5]

$$\left. \begin{aligned} |\lambda|\ L' &= |\lambda - 1|\ L \\ \text{and} \qquad & \\ L^2 &= \Omega^2 \lambda^2 / 4\beta \end{aligned} \right\} \tag{12}$$

From (2) and (3) we have

$$\mathrm{d}\phi/\mathrm{d}z = \Omega\,(\lambda - z)/2z\,(1 - z)\ [H(z)]^{1/2} = g(z). \tag{13}$$

Hence the apsidal angle Φ may now be written as

$$\Phi = \int_{z_2}^{z_3} g(z)\,\mathrm{d}z. \tag{14}$$

Substituting, (see Rouch and Lebowitz, [8])

$$z' = (z - z_3)\ (z_4 - z_2)/(z - z_2)\ (z_4 - z_3), \tag{15}$$

(14) may be written (writing z for z') as

$$\begin{aligned} \Phi &= 2^{-1}\Omega\ [4\beta\,(z_1 - z_2)\ (z_4 - z_3)\,]^{-1/2} \\ &\times \int_0^\infty [\lambda/(zz_2 - z_3K) - (1 - \lambda)/\{z\,(1 - z_2) + (z_3 - 1)K\}] \\ &\times [H_1\,(z)]^{-1/2}\ (z - K)\,\mathrm{d}z, \end{aligned} \tag{16}$$

where

$$H_1\,(z) = z\,(1 - z)\ (\tilde{\lambda}^{-1} - z), \tag{17}$$

$$\tilde{\lambda}^{-1} = K\,(z_1 - z_3)/(z_1 - z_2) \tag{18}$$

and

$$K = (z_4 - z_2)/(z_4 - z_3), \tag{19}$$

Again as [5], we can write

$$\Phi = \mathrm{sgn}\ (\lambda)\ \Phi_1 + \mathrm{sgn}\ (\lambda - 1)\ \Phi_2 \tag{20}$$

where we have straight away

$$\Phi_1 = \int_{z_2}^{z_3} g_1\,(z)\,\mathrm{d}z = \int_{z_2}^{z_3} L\ \mathrm{d}z/2z\ [H(z)]^{1/2} \tag{21}$$

and

$$\Phi_2 = \int_{z_2}^{z_3} g_2\,(z)\,\mathrm{d}z = \int_{z_2}^{z_3} L'\ \mathrm{d}z/2\ (1 - z)\ [H(z)]^{1/2}. \tag{22}$$

Again under the transformation (15) Φ may be given by

$$\Phi_i = 2^{-1}[(z_1 - z_2)(z_4 - z_3)]^{-1/2} \times \int_0^\infty G_i(z)\,dz, \quad (i = 1, 2) \tag{23}$$

where

$$G_1(z) = L(z - K)/(zz_2 - z_3 K)\;[H_1(z)]^{1/2} \tag{24}$$

and

$$G_2(z) = L'(z - K)/[(1 - z)\,z - (1 - z_3)K] \times [H_1(z)]^{1/2} \tag{25}$$

3. Halphen's inequality

To establish the lower bounds for Φ, we need to prove the following.

Proposition 1

$$\Phi_1 = \pi/2 + P_1, \tag{26}$$

and

$$\Phi_2 = -\pi/2 + P_2, \tag{27}$$

where $P_i\,(i = 1, 2)$ are two positive integrals and $\Phi_i\,(i = 1, 2)$ are as given in (23).

We shall use the method of complex integration to prove the above proposition. Considering the integrand in (23) as a function of the complex variable, we obtain a single valued branch of this function on the Riemann sheets bounded by the cuts $C_j\,(j = 1, 2)$ and the circle C_0 with its centre at the origin (see figure 1). By putting

$$z - z_k = r_k \exp(i\,\theta_k), \quad (0 \leqslant \theta_k < 2\pi), \quad (k = 1, 2, 3). \tag{28}$$

The signs of $[H_1(z)]^{1/2}$ on the cuts have been fixed and shown in figure 1.

It may be noted that the signs of $(z - K)$ and $(zz_2 - z_3K)$ are negative for $0 < z < 1$ and those of $(z - k)$ and $[(1 - z_2)\,z - (1 - z_3)K]$ are positive for $z > \tilde{\lambda}^{-1}$. Integrating around the contour $C_0 - C_1 - C_2$ and applying Cauchy's residue theorem we have

$$\int_{C_0} G_i(z)dz - \int_{C_1} - \int_{C_2} = 2\pi i R, \tag{29}$$

where R represents the sum of the residues of the integrand concerned. As usual $\int_{C_0} \to 0$ as the circle is made infinitely large. The contributions due to the other remaining integrals are

$$\int_{C_1} = -\int_0^1 G_1(z)\,dz, \tag{30}$$

$$\int_{C_2} = \int_{\tilde{\lambda}^{-1}}^\infty G_1(z)\,dz, \tag{31}$$

Similarly for $G_2(z)$, we have

$$\int_{C_1} = -\int_0^1 G_2(z)\,dz, \tag{32}$$

$$\int_{C_2} = \int_{\lambda^{-1}}^{\infty} G_2(z)\,dz. \tag{33}$$

The residues of $G_1(z)$ at $z = z_3K/z_2$ and $G_2(z)$ at $z = (1 - z_3)K/(1 - z_2)$ are $1/2i$ and $-1/2i$ respectively. Now using integrals (30) to (33) in (29) we have (26) and (27) with

$$P_1 = (1/2)\int_1^{\infty} G_1(z)\,dz + (1/2)\int_{\lambda^{-1}}^{\infty} G_1(z)\,dz > 0 \tag{34}$$

and

$$P_2 = (1/2)\int_1^{\infty} G_2(z)\,dz + (1/2)\int_{\lambda^{-1}}^{\infty} G_2(z)\,dz > 0. \tag{35}$$

Thus Proposition-1 is proved.

With the help of the above proposition we shall now prove the following.

Theorem 1

$$\Phi < -\pi/2 \text{ when } \lambda \leqslant 0, \tag{36}$$

and

$$\Phi > \pi/2 \text{ when } \lambda \geqslant 1. \tag{37}$$

Proof

When $\lambda < 0$, we have

$$\begin{aligned} \Phi &= -(\Phi_1 + \Phi_2), && \text{(due to (20))} \\ &= -\pi/2 - (P_1 + \Phi_2), && \text{(due to (26))} \\ &< -\pi/2. && \text{(since } \Phi_2, P_1 > 0) \end{aligned} \tag{38}$$

When $\lambda = 0$, the precessional advance defined by the integral $\int_0^{\infty} G(z)\,dz$ degenerates to

$$\Phi = 2^{-1}[4\beta(z_1 - z_2)(z_4 - z_3)]^{-1/2} \times \int_0^{\infty}(z - K)\,dz/[(1 - z_2)z - (1 - z_3)K] \times (H_1(z))^{1/2}. \tag{39}$$

Evaluating the integral contained in (39) round the contour $C_0 - C_1 - C_2$ (figure 1) and using (10), (11) and (13) we have

$$\Phi = -\pi/2 - 1/2\left[\int_1^{\infty} G(z)\,dz + \int_{\lambda^{-1}}^{\infty} G(z)\,dz\right] \tag{40}$$

$< -\pi/2$, Since the integrals contained in parenthesis above are positive.

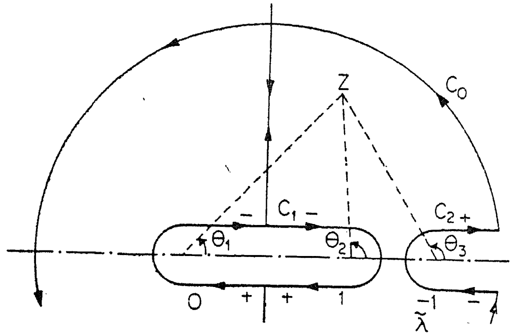


Figure 1. Riemann sheets for $Gi(z)$.

Now (38) and (40) together prove the first part of the theorem.

Coming to the second part of the theorem we shall consider first the case when $\lambda = 1$.

In this case

$$\Phi = \Phi_1 = \pi/2 + P_1 \qquad \text{(due to Proposition 1)}$$

$$> \pi/2. \qquad \text{(as } P_1 > 0) \tag{41}$$

Again for $\lambda > 1$, we have

$$\Phi = \Phi_1 + \Phi_2$$

$$= \pi/2 + (P_1 + \Phi_2)$$

$$> \pi/2. \qquad \text{(as usual)} \tag{42}$$

Thus due to (41) and (42), the second part of the theorem is proved. In this way the Halphen's inequality is demonstrated. The same inequality has already been proved by different method (*cf* Rath and Pal [7]).

4. Sequences of motion and the limiting quartic

Here we shall construct two sequences of motion as Diaz and Metcalf [2], (also see [3]) to establish the upper bounds for Φ. These are, when $t = 0$,

$$(a) \quad z_n(0) = 1/2, \quad z_n'(0) = 0, \quad \phi_n(0) = 0$$

and

$$(b) \quad z_n(0) = 1/2, \quad z_n'(0) = \gamma, \quad \phi_n(0) = 0.$$

with the stipulation that

$$q_n, \ \Omega_n \to 0 \text{ as } n \to \infty$$

such that

$$F_n = (2\lambda - 1)\,\Omega_n, \quad \alpha_n + 2\beta_n = 2\mu(0)/B = C, \text{ say}$$

where λ and C are independent of n, γ in (*b*) is chosen to be an arbitrarily large positive constant, which is also independent of n.

Now in a straight forward manner (*cf* [2], also [3]), the limiting quartics for the two cases (*a*) and (*b*) can found out to be

$$\underset{n\to\infty}{Lt}\ H_n(z) = Cz(1-2z)(1-z)$$

so that the zeros of $H_n(z)$ in the limit are $-\infty$, 0, 1/2 and 1, and

$$\underset{n\to\infty}{Lt}\ H_n(z) = z(1-z)(4^2\gamma + C - 2Cz)$$

Thus the limiting values of the zeros of $H_n(z)$ are

$$-\infty,\ 0,\ 1 \text{ and } (4^2\gamma + C)/2C$$

The last root exceed 1, as the arbitrary γ is chosen such that it exceeds $(C)^{1/2}/2$.

For the sequence of motion (*a*), equation (7) implies that

$$E_1 \to 0 \text{ and hence } H_n(1/2) \to 0.$$

Consequently, it can be shown that

$$H_n(z) = z(1-z)(2z-1)(\beta_n - \alpha_n - 2\beta_n z),$$

which gives the zeros of the quartic as

$$\left.\begin{array}{l} z_{1n} = (\beta_n - \alpha_n)/\beta_n = -K_{1n} \quad \text{(say)}, \\ z_{2n} = 0,\ z_{3n} = 1/2,\ z_{4n} = 1. \end{array}\right\}$$

Here K_{1n} is positive and sufficiently large.

Likewise, for the sequence, (b)

$$E_{1n} \to 4\gamma^2$$

and

$$H_n(z) = z(1-z)(z + K_{2n})(1 - K_{3n})$$

where

$$K_{2n} = [2\beta_n - \alpha_n - (\alpha_n^2 + 16\beta_n\gamma^2)^{1/2}]/2\beta_n$$

and

$$K_{3n} = (2\beta_n - \alpha_n + (\alpha_n^2 + 16\beta_n\gamma^2)^{1/2}]/2\beta.$$

Here K_{2n} and K_{3n} are positive and sufficiently large, as γ has been chosen sufficiently large. It may be noted that on the contrary if $\gamma \to 0$, then this case becomes the same as (*a*). Thus, the zeros of the quartic $H_n(z)$ are given by

$$z_{1n} = -K_{2n},\quad z_{2n} = 0,\quad z_{3n} = 1,\quad z_{4n} = K_{3n} \tag{44}$$

5. Limiting Upper Bounds of Φ

To establish the upper bound of Φ, we shall use the following result, due to (Rath and Pal)

$$\Phi = -\pi - Q \quad \text{for} \quad z_1 \leqslant \lambda < 0, \tag{45}$$

and

$$\Phi = \pi + Q \quad \text{for} \quad 1 < \lambda \leqslant z_4 \tag{46}$$

where

$$Q = -\int_{-\infty}^{z_1} + \int_{z_4}^{\infty} \Omega\,(\lambda - z)/2z\,(1 - z)\,[\prod_{i=1}^{4}(z - z_i)]^{1/2}$$

which can be rewritten as in accordance with the sequence of motion (b) as

$$Q_n = -\int_{-\infty}^{-K_{2n}} + \int_{K_{3n}}^{\infty} \frac{\Omega_n\,(\lambda - z)\,dz}{2z\,(1 - z)\,[(z + K_{2n})\,(z)\,(z - 1)\,(z - K_{3n})]^{1/2}} \tag{47}$$

Setting $\Phi + \pi = \sigma$ (48)

We have from (45) and (47)

$$\frac{\sigma}{\Omega_n} = -\int_{-\infty}^{-K_{2n}} + \int_{K_{3n}}^{\infty} \frac{(\lambda - z)\,dz}{2z\,(1 - z)\,[(z + K_{2n})\,(z)\,(z - 1)\,(z - K_{3n})]^{1/2}} \tag{49}$$

Since the integrand is finite over the respective ranges of integration on the right hand side of (49), $\sigma \to 0$ with $\Omega_n \to 0$, in that case $\Phi \to -\pi$ when $\lambda < 0$ and similarly it can be proved from (46) and (47) that $\Phi \to \pi$. Hence Puiseux inequality is established along with its exactness.

6. On the exactness of Halphen's bounds,

To prove the exactness of Halphen's bounds, we shall consider the following cases using the sequence of motion (a) given in the preceding sections.

Case 1 $\lambda < 0$.

It is known that

$$\Phi < -\pi/2, \text{ when } \lambda < 0,$$

we shall show that $-\pi/2$ is the lowest upper bound for Φ.

Now, corresponding to the sequence of motion (a) for which

$$z_{1n} = -K_{1n},\ K_{2n} = 0,\ z_{3n} = 1/2,\ z_{4n} = 1,$$

we have from (20) and using equations (10 to 14)

$$\Phi_n = -1/2\,(z_{2n} \times z_{3n})^{1/2} \left[\int_{z_{2n}}^{1/2} \frac{(-z_{1n}z_{4n})^{1/2}\, dz}{z\,[\prod_{i=1}^{4} (z - z_{in})]^{1/2}} \right.$$

$$\left. + \frac{|\lambda - 1|}{|\lambda|} \int_{z_{2n}}^{1/2} \frac{(-z_{1n}z_{4n})^{\frac{1}{2}} dz}{(1-z)\,[\prod_{i=1}^{4} (z - z_{in})]^{1/2}} \right] \tag{50}$$

Defining certain auxiliary integrals (Rath and Namboodiri [5]), as

$$I_1 = 2^{-1} \int_{z_2}^{z_3} [z_2 z_3/(z - z_2)\,(z_3 - z)]^{1/2}\, dz/z, \tag{51}$$

$$I_2 = 2^{-1} \int_{z_2}^{z_3} [(1 - z_2)\,(1 - z_3)/(z - z_2)\,(z_3 - z)]^{1/2}\, dz/(1 - z) \tag{52}$$

$$I_3 = 2^{-1} \int_{z_2}^{z_3} [(z - z_2)\,(z_3 - z)]^{-1/2}\, dz \tag{53}$$

where z_2 and z_3 are such that $0 < z_2 < z_3 < 1$, and $I_1 = I_2 = I_3 = \pi/2$.

It follows from (50) and (51) that

$$-\pi/2 - \Phi_n = (z_{2n})^{1/2} \times 2^{-3/2} \left[J_{1n} + \frac{|\lambda - 1|}{|\lambda|} J_{2n} \right], \text{ say} \tag{54}$$

where

$$J_{1n} = \int_{z_{2n}}^{1/2} \frac{[(-z_{1n}z_{4n})^{1/2} - \{(z - z_{1n})\,(z_{4n} - z)\}^{1/2}]\, dz}{z\,[\prod_{i=1}^{4} (z - z_{in})]^{1/2}}$$

$$\leqslant 2C_1 I_3 \text{ (due to first mean-value theorem, } cf \text{ [1])}$$

$$= C_1 \pi, \text{ (due to 53)} \tag{55}$$

($C_1 > 0$, is a constant independent of z and n) and as before

$$J_{2n} = \int_{z_{2n}}^{1/2} \frac{(-z_{1n}z_{4n})^{1/2}\, dz}{(1 - z)\,[\prod_{i=1}^{4} (z - z_{in})]^{1/2}}$$

$$\leqslant 2C_2 I_2 \text{ (first mean value theorem)}$$

$$= \pi\, C_2, \text{ (due to 52)} \tag{56}$$

($C_2 > 0$, is a constant independent of z and n).

Using (55) and (56) in (54), we have

$$-\Phi_n - \pi/2 < (z_{2n})^{1/2} \times 2^{-3/2} \left[C_1 + \frac{|\lambda - 1|}{|\lambda|} C_2 \right] \pi$$

since $z_{2n} = 0$, (see 43).

Clearly $\Phi_n \to -\pi/2$, which demonstrates the sharpness of Halphen's bound.

Case 2 $\lambda > 1$.

As in the preceding case, it may now be shown in a straight forward manne that

$$\Phi_n - \pi/2 \to 0 \text{ as } n \to \infty$$

which is the result required.

Case 3 $\lambda = 0$,

We have from (20)

$$\Phi_n = -\Phi_{2n} = -\frac{1}{2}\int_{z_{2n}}^{1/2} \frac{\left|\prod_{i=1}^{4}(1 - z_{in})\right|^{1/2} dz}{(1-z)\left\{\prod_{i=1}^{4}(z - z_{in})\right\}^{1/2}}$$

$$= -\frac{1}{2}\int_{z_{2n}}^{1/2} \frac{\{(1-z_{2n})(1-z_{3n})\}^{1/2}}{(1-z)\{(z-z_{2n})(z_{3n}-z)\}^{1/2}} \times \frac{\{(1-z_{1n})(z_{4n}-1)\}^{1/2}}{\{(z-z_{1n})(z_{4n}-z)\}^{1/2}}\, d$$

$$\leqslant -C_n\, \pi/2 \qquad (57$$

where,

$$C_n = \max\left\{\frac{(1-z_{1n})(z_{4n}-1)}{(z-z_{1n})(z_{4n}-z)}\right\}^{1/2}$$

$$= \left\{\frac{(1-z_{1n})(z_{4n}-1)}{-z_{1n}(z_{4n}-1)}\right\}^{1/2}, \text{ due to (9)}$$

$$\sim \left(\frac{1+k_{1n}}{k_{1n}}\right), \text{ due to (43)}$$

In the limit, when $n \to \infty$, $k_{1_n} \to \infty$, $C_n \to 1$ and hence $\Phi_n \to -\pi/2$, whicl proves the sharpness of the bound.

Case 4, $\lambda = 1$

As before from (20) we have

$$\Phi_n = \frac{1}{2}\int_{z_{2n}}^{1/2} \frac{\left[-\prod_{i=1}^{4} z_{in}\right]^{1/2} dz}{z\left[\prod_{i=1}^{4}(z - z_{in})\right]^{1/2}}$$

which can be rewritten using (51) as

$$\Phi_n - \frac{\pi}{2} = (z_{2n})^{1/2} \times 2^{-3/2}\int_{z_{2n}}^{1/2} \frac{[(-z_{1n} z_{4n})^{1/2} - \{(z - z_{1n})(z_{4n} - z)\}^{1/2}]\, dz}{z\left[\prod_{i=1}^{4}(z - z_{in})\right]^{1/2}}$$

$$\leqslant (z_{2n})^{1/2}\, C_4\, I_3, \text{ (first mean value theorem)}$$

$$= (z_{2n})^{1/2}\, C_4\, \pi, \text{ (due to 52)} \qquad (58)$$

($C_4 > 0$, a constant independent of z and n).

Since $z_{2n} = 0$, (see (43))
(58) shows that $\Phi_n - \pi/2 \to 0$, which now demonstrates the sharpness of the required bound.

7. Conclusion

The Halphen Puiseux limits of the librations which we have discussed in the preceding section clearly pertain to an aerodynamic gyroscope. This becomes a Lagrangian gyroscope if one stipulates that aerodynamic term $q = 0$ and $\mu(0)$ which represents the distance of the aerodynamic centre of pressure from the centre of mass of the missile, measured along the axis of dynamic and aerodynamic symmetry. In analogy with a gravity gyro only one has to set $\mu(0) = \text{mg a}$ (*cf* [9]).

List of symbols

A	The axial moments of inertia of the projectile
B	The transverse moment of inertia of the projectile
C_o	The circle with its centre at the origin
$C_j\,(j = 1, 2)$	The cuts on the Reimann sheets
C_1, C_2, C_n, C_4	The constants, independent of z and n
E	The total constant energy
F	The constant of angular momentum
I_1, I_2, I_3	Three positive integrals each equal to $\pi/2$
K	The ratio of the roots of the quartic, $\left(\frac{z_4 - z_2}{z_4 - z_3}\right)$
p_1, p_2, Q	The positive integrals
q	Certain aerodynamic dimensionless parameters
R	Sum of the residues of the integrand concerned
N	The constant axial spin
$z_i\,(i = 1 \text{ to } 4)$	The zeros of quartic, $H(z)$
$\mu(0)$	Certain aerodynamic parameter
γ	An arbitrarily large positive constant
α, β	The constant components of angular momentum about the vertical and the axes of symmetry
ϕ	The angle of precession
Φ	The precessional advance
δ	Angle of nutation
λ	Ratio of angular momentum
Ω	Spin parameter

References

[1] Brand L 1955 *Advanced calculus* (New York : John Wiley and Sons, Inc) p. 266
[2] Diaz J B and Metcalf F T 1964 Upper and lower bounds for the apsidal angle in the theory of the heavy symmetrical top, *Arch. Rat. Mech. Anal.* **16** 214–229

[3] Kohn W 1946 Contour integration in the theory of the spherical pendulum and the heavy symmetrical top, *Trans. Am. Math. Soc.* **59** 107–131

[4] Rath P C and Namboodiri A V 1974 On the Lock-Fowler model of a spinning shell *Indian J. Pure Appl. Math.* **5** 396–418

[5] Rath P C and Namboodiri A V 1978 The apsidal limits of a rolling missile ; *Q. J. Appl. Math.* **36** 1–17

[6] Rath P C and Namboodiri AV 1980 Librations of a Lock-Fowler projectile, *Memoria de L'Artillerie Francaise* **54** 2FASC, 313–349

[7] Rath P C and Pal J Precessional advances for a Lock-Fowler missile ; *Q.J. Appl. Math.* **39** 375–381

[8] Rouch H E and Lebowitz A 1973 *Elliptic functions, theta functions and Riemann surfaces,* Baltimore, Maryland, p. 8

[9] Synge J L and Friffith B A 1959 Principles of mechanics (New York : McGraw-Hill Book Co. Inc.) p. 393

PROCEEDINGS OF THE

INDIAN ACADEMY OF SCIENCES

(Mathematical Sciences)

VOLUME 91, JANUARY-DECEMBER 1982

INDEX

THE INDIAN ACADEMY OF SCIENCES
BANGALORE 560080

PROCEEDINGS (Mathematical Sciences)

Vol. 91, January–December 1982

SUBJECT INDEX (Mathematical Sciences)

AUTHOR INDEX (Mathematical Sciences)